

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**





101 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTGCGCGGT 147  
|||||  
34 rgrprogllyserglyserclnlsprtsrtyrlyleuargprogly 49

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-694-865-4

seq\_documentation\_block:

; Sequence 4, Application US/08694865  
; Patent No. 5837268

; GENERAL INFORMATION:

; APPLICANT: POTTER, ANDREW A.

; APPLICANT: MANN, JOHN G.

; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: REED & ROBINS LLP

; STREET: 285 HAMILTON AVENUE, SUITE 200

; CITY: PALO ALTO

; STATE: CA

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/694,865

; FILING DATE: 09-AUG-1996

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: MCCracken, THOMAS P.

; REGISTRATION NUMBER: 38,548

; REFERENCE/DOCKET NUMBER: 9001-0016.22

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415)327-3400

; TELEFAX: (415)327-3231

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 49 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-694-865-4

alignment\_scores:

Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x US-08-694-865-4 ..

Align seg 1/1 to: US-08-694-865-4 from: 1 to: 49

1 CAGATTGAGAGCTAGCGCTGCGCGGTCTCAGATTGAG 50  
|||||  
1 GlnIstPrsErTyRgLyLeuArgProGlySerGlnIstPrsE 17  
51 CTAGGCGCTGCGGTCCGGTGGCTCTAGCCAGCATTTGAGCTAGCGCGCTGC 100  
|||||  
17 rTyGlyLeuArgProGlySerSerGlnIstPrsErTyRgLyLeuA 34  
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTGCGCGGT 147  
|||||  
34 rgrprogllyserglyserclnlsprtsrtyrlyleuargprogly 49

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-878-748-4

seq\_documentation\_block:

; Sequence 4, Application US/08878748

; Patent No. 5969126

; GENERAL INFORMATION:  
; APPLICANT: POTTER, ANDREW A.  
; APPLICANT: REDMOND, MARK J.  
; APPLICANT: HUGHES, HOW P.A.  
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: REED & ROBINS

; STREET: 635 BRYANT STREET

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/878,748

; FILING DATE: 19-JUN-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/387,156

; FILING DATE: 10-FEB-1995

; APPLICATION NUMBER: US 07/960,932

; FILING DATE: 14-OCT-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/779,171

; FILING DATE: 16-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: ROBINS, ROBERTA L.

; REGISTRATION NUMBER: 33,208

; REFERENCE/DOCKET NUMBER: 9001-0016.21

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 617-8999

; TELEFAX: (415) 327-3231

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 49 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-878-748-4

alignment\_scores:

Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x US-08-878-748-4 ..

Align seg 1/1 to: US-08-878-748-4 from: 1 to: 49

1 CAGATTGAGAGCTAGCGCTGCGCGGTCTCAGATTGAG 50  
|||||  
1 GlnIstPrsErTyRgLyLeuArgProGlySerGlnIstPrsE 17  
51 CTAGGCGCTGCGGTCCGGTGGCTCTAGCCAGCATTTGAGCTAGCGCGCTGC 100  
|||||  
17 rTyGlyLeuArgProGlySerSerGlnIstPrsErTyRgLyLeuA 34  
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTGCGCGGT 147  
|||||  
34 rgrprogllyserglyserclnlsprtsrtyrlyleuargprogly 49

seq\_name: /cgn2\_6/ptodata/2/1aa/6\_COMB.pep:US-09-124-491-4

seq\_documentation\_block:

; Sequence 4, Application US/09124491

; Patent No. 6022960

```

: GENERAL INFORMATION:
: APPLICANT: POTTER, ANDREW A.
: APPLICANT: MANN, JOHN G.
: TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: REED & ROBINS LLP
: STREET: 285 HAMILTON AVENUE, SUITE 200
: CITY: PALO ALTO
: STATE: CA
: COUNTRY: USA
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/124,491
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/694,865
: FILING DATE: 09-AUG-1996
: APPLICATION NUMBER: US 08/387,156
: FILING DATE: 10-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/960,932
: FILING DATE: 14-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/779,171
: FILING DATE: 16-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: MCCracken, THOMAS P.
: REGISTRATION NUMBER: 38,548
: REFERENCE/DOCKET NUMBER: 9001-0016.22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)327-3400
: TELEFAX: (415)327-3231
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-124-491-4

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-09-124-491-4  ..
Align seg 1/1 to: US-09-124-491-4 from: 1 to: 49
1 CAGCATTTGAGCTACGGCTCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnH1StrpSerTyrGlyLeuArGProGlySerGlySerGlnAspTrpSe 17
51 CTACGGCCCTCGGCTCGGCTGCTTACGACGATTGGAGCTACGGCCGTC 100
|||||
17 rTyGrIyLeuArGProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 34
101 GCCCTGGCAGCGGTAGCCCAAGATTGGAGCTACGGCTCGCCGGGT 147
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArGProGly 49
seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-387-156-10
seq_documentation_block:

```

```

: Sequence 10, Application US/08387156
: Patent No. 5723129
: GENERAL INFORMATION:
: APPLICANT: POTTER, ANDREW A.
: APPLICANT: REDMOND, MARK J.
: APPLICANT: HUGHES, HOW P. A.
: TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: REED & ROBINS
: STREET: 635 BRYANT STREET
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/387,156
: FILING DATE: 10-FEB-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/960,932
: FILING DATE: 14-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/779,171
: FILING DATE: 16-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: ROBINS, ROBERTA L.
: REGISTRATION NUMBER: 33,208
: REFERENCE/DOCKET NUMBER: 9001-0016.21
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 617-8999
: TELEFAX: (415) 327-3231
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 544 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-387-156-10

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-387-156-10  ..
Align seg 1/1 to: US-08-387-156-10 from: 1 to: 544
1 CAGCATTTGAGCTACGGCTCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
494 GlnH1StrpSerTyrGlyLeuArGProGlySerGlySerGlnAspTrpSe 510
51 CTACGGCCCTCGGCTCGGCTGCTTACGACGATTGGAGCTACGGCCGTC 100
|||||
510 rTyGrIyLeuArGProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 527
101 GCCCTGGCAGCGGTAGCCCAAGATTGGAGCTACGGCTCGCCGGGT 147
|||||
527 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArGProGly 542
seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-694-865-10
seq_documentation_block:
: Sequence 10, Application US/08694865
: Patent No. 5837268

```

GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,865  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCRAKEN, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-865-10

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x US-08-694-865-10 ..  
Align seg 1/1 to: US-08-694-865-10 from: 1 to: 544

1 CAGCATTGAGACTAGGCGCTGCGCCCTGCGCAGCGGTTCTCAAGATTGGAG 50  
|||||  
494 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 510  
51 CTAGCGGCTGCGTCCGGGTGCTACGACATGGAGCTAGCGGCTGC 100  
|||||  
510 rTyGlyLeuArgProGlySerGlySerGlnHisTrpSerTyrGlyLeuA 527  
101 GCCCTGACAGCGGTAGCCAGATTGGAGCTAGCGCTGCGGGT 147  
|||||  
527 rGrProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 542

seq\_name: /cgn2.6/ptodata/2/1aa/5B\_COMB.pep:US-08-878-748-10

seq\_documentation\_block:  
Sequence 10, Application US/08878748  
Patent No. 5969126  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P.A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET

CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/878,748  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-878-748-10

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x US-08-878-748-10 ..  
Align seg 1/1 to: US-08-878-748-10 from: 1 to: 544

1 CAGCATTGAGACTAGGCGCTGCGCCCTGCGCAGCGGTTCTCAAGATTGGAG 50  
|||||  
494 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 510  
51 CTAGCGGCTGCGTCCGGGTGCTACGACATGGAGCTAGCGGCTGC 100  
|||||  
510 rTyGlyLeuArgProGlySerGlySerGlnHisTrpSerTyrGlyLeuA 527  
101 GCCCTGACAGCGGTAGCCAGATTGGAGCTAGCGCTGCGGGT 147  
|||||  
527 rGrProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 542

seq\_name: /cgn2.6/ptodata/2/1aa/6\_COMB.pep:US-09-124-491-10

seq\_documentation\_block:  
Sequence 10, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO

STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-10

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x US-09-124-491-10 ..  
Align seg 1/1 to: US-09-124-491-10 from: 1 to: 544

1 CAGCATGTGAGCTACGGCCCTGCGCCCTGCGACGGTTCTCAAGATTGAG 50  
|||||  
494 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeu 510  
51 CTACGGCGCTGCGGTGCGGCTGCTAGCCAGATTGAGCTAGCGCTCGCGGT 100  
|||||  
510 TyrGlyLeuArgProGlyGlySerSerGlnHisTrpSerTyrGlyLeu 527  
527 rgProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 542

seq\_name: /cgn2\_6/ptodata/2/1aa/58\_COMB.pep:US-08-694-865-16

seq\_documentation\_block:  
Sequence 16, Application US/08694865  
Patent No. 5837268  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP

STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,865  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 699 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-865-16

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x US-08-694-865-16 ..  
Align seg 1/1 to: US-08-694-865-16 from: 1 to: 699

1 CAGCATGTGAGCTACGGCCCTGCGCCCTGCGACGGTTCTCAAGATTGAG 50  
|||||  
9 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeu 25  
25 TyrGlyLeuArgProGlyGlySerSerGlnHisTrpSerTyrGlyLeu 42  
42 rgProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 57

seq\_name: /cgn2\_6/ptodata/2/1aa/6\_COMB.pep:US-09-124-491-16

seq\_documentation\_block:  
Sequence 16, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, Thomas P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 699 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-16

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x US-09-124-491-16 ..

Align seq 1/1 to: US-09-124-491-16 from: 1 to: 699

```
1 CAGCATTGAGCTAGCGCGCTGCGCGGCTTCTCAAGATTGAG 50
|||||
9 GlnIstPrSerTyrGlyLeuArgProGlySerGlnAspTrpSe 25
51 CTAGGCGCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTAGCGCGCTGC 100
|||||
25 rTyGlyLeuArgProGlySerGlnIstPrSerTyrGlyLeuA 42
101 GCCCTGCGAGCGGTAGCCAGATTGAGCTAGCGCTGCGCGGT 147
|||||
42 rGrProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 57
```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-387-156-8

seq\_documentation\_block:

Sequence 8, Application US/08387156  
Patent No. 5723129  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HUW P.A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/387,156  
FILING DATE: 10-FEB-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-387-156-8

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x US-08-387-156-8 ..

Align seq 1/1 to: US-08-387-156-8 from: 1 to: 977

```
1 CAGCATTGAGCTAGCGCGCTGCGCGGCTTCTCAAGATTGAG 50
|||||
927 GlnIstPrSerTyrGlyLeuArgProGlySerGlnAspTrpSe 943
51 CTAGGCGCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTAGCGCGCTGC 100
|||||
943 rTyGlyLeuArgProGlySerGlnIstPrSerTyrGlyLeuA 960
101 GCCCTGCGAGCGGTAGCCAGATTGAGCTAGCGCTGCGCGGT 147
|||||
960 rGrProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 975
```

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-694-865-8

seq\_documentation\_block:

Sequence 8, Application US/08694865  
Patent No. 5837268  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/694,865
  FILING DATE: 09-AUG-1996
  CLASSIFICATION: 424
  ATTORNEY/AGENT INFORMATION:
    NAME: MCCracken, THOMAS P.
    REGISTRATION NUMBER: 38,548
  REFERENCE/DOCKET NUMBER: 9001-0016.22
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (415)327-3400
    TELEFAX: (415)327-3231
  INFORMATION FOR SEQ ID NO: 8:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 977 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-694-865-8

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-694-865-8 ..

Align seg 1/1 to: US-08-694-865-8 from: 1 to: 977

1 CAGCATTTGAGCTAGCGCTCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
927 GlnHstPserTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 943
943 rTyrGlyLeuArgProGlyGlySerSerGlnHstPserTyrGlyLeuA 960
51 CTAAGCGCTGCGTCCGGGTGCTAGCCAGCATTTGAGCTAGCGGCTGC 100
|||||
943 rTyrGlyLeuArgProGlyGlySerSerGlnHstPserTyrGlyLeuA 960
101 GCCCTGGAGCGCGTAGCCAGATTGGAGCTAGCGCTCGCTCGGGGT 147
|||||
960 rGProGlySerGlySerGlnAspTrpserTyrGlyLeuArgProGly 975

seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-878-748-8

seq_documentation_block:
; Sequence 8, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P. A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
```

```
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/779,171
  FILING DATE: 16-OCT-1991
  ATTORNEY/AGENT INFORMATION:
    NAME: ROBINS, ROBERTA L.
    REGISTRATION NUMBER: 33,208
  REFERENCE/DOCKET NUMBER: 9001-0016.21
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (415) 617-8999
    TELEFAX: (415) 327-3231
  INFORMATION FOR SEQ ID NO: 8:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 977 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-878-748-8

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-878-748-8 ..

Align seg 1/1 to: US-08-878-748-8 from: 1 to: 977

1 CAGCATTTGAGCTAGCGCTCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
927 GlnHstPserTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 943
943 rTyrGlyLeuArgProGlyGlySerSerGlnHstPserTyrGlyLeuA 960
51 CTAAGCGCTGCGTCCGGGTGCTAGCCAGCATTTGAGCTAGCGGCTGC 100
|||||
943 rTyrGlyLeuArgProGlyGlySerSerGlnHstPserTyrGlyLeuA 960
101 GCCCTGGAGCGCGTAGCCAGATTGGAGCTAGCGCTCGCTCGGGGT 147
|||||
960 rGProGlySerGlySerGlnAspTrpserTyrGlyLeuArgProGly 975

seq_name: /cgn2_6/ptodata/2/iaa/6_COMB.pep:US-09-124-491-8

seq_documentation_block:
; Sequence 8, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE: 09-AUG-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
```

FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-8

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x US-09-124-491-8 ..

Align seg 1/1 to: US-09-124-491-8 from: 1 to: 977

1 CAGATTGAGCTAGCGGCTGCGCCGCGAGCGGCTCTCAAGATTGAG 50  
|||||  
927 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnAspTrpSe 943  
51 CTAGGCGCTGCTCGGCTGCTAGCCAGCATTTGAGCTAGCGCTGC 100  
|||||  
943 rTyGlyLeuArgProGlySerSerGlnHisTrpSerTyrGlyLeu 960  
101 GCCCTGCGAGCGGTAGCCAGATTGAGCTAGCGCTGCGCGGT 147  
|||||  
960 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 975

seq\_name: /cgn2\_6/plodata/2/1aa/5A\_COMB.pep:US-07-690-983D-45

seq\_documentation\_block:

Sequence 45, Application US/07690983D  
Patent No. 5403586  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J.  
APPLICANT: STEWART, Andrew G.  
APPLICANT: TSONIS, Con G.  
TITLE OF INVENTION: FUSION PROTEINS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/690,983D  
FILING DATE: 25-JUN-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU90/00373

FILING DATE: 24-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/148 CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 44 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-690-983D-45

alignment\_scores:  
Quality: 187.50 Length: 49  
Ratio: 4.934 Gaps: 3  
Percent Similarity: 77.551 Percent Identity: 69.388

alignment\_block:  
US-09-306-689-10 x US-07-690-983D-45 ..

Align seg 1/1 to: US-07-690-983D-45 from: 1 to: 44

1 CAGATTGAGCTAGCGGCTGCGCCCTGCGAGCGGCTCTCAAGATTGAG 50  
:::|||||  
3 GlnHisTrpSerTyrGlyLeuArgProGly.....GlnHisTrpSe 16  
51 CTAGGCGCTGCTCGGCTGCTAGCCAGCATTTGAGCTAGCGCTGC 100  
|||||  
16 rTyGlyLeuArgProGly.....GlnHisTrpSerTyrGlyLeu 30  
101 GCCCTGCGAGCGGTAGCCAGATTGAGCTAGCGCTGCGCGGT 147  
|||||  
30 rGProGly.....GlnHisTrpSerTyrGlyLeuArgProGly 42



OM of: US-09-306-689-10 to: SPTREMBL\_15:\* out\_format : pfs

Date: Mar 2, 2001 10:49 AM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

#### Command line parameters:

-MODEL=frimex-n2p.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool/US09306689/runat.02032001\_102828.9616/app\_query.fasta\_1.2389  
-DB=SPTREMBL\_15 -QFMT=fastan -SUFFIX=rspt -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DEEXT=7.000 -START=1 -MATRIX=blomsun62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR=SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -MINLEN=0 -MAXLEN=2000000000  
-USER=US09306689.etcgnl\_1.639 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
-MAIT -THREADS=1

#### Search information block:

Query: US-09-306-689-10  
Query length: 147  
Database: SPTREMBL\_15:\*  
Database sequences: 374700  
Database length: 117207915  
Search time (sec): 337.480000

#### score\_list:

Sequence	Strd	Orig	ZScore	Escore	len	Documentation
sp_vertibrate:09PT52	+	95.50	166.91	0.0033	263	09PT52 agkistrodon halys blomh
sp_invertebrate:09WTY8	+	94.50	166.92	0.0037	836	09WTY8 rattus norvegicus (rat)
sp_invertebrate:09U617	+	81.00	153.65	0.1283	1729	09U617 drosophila melanogast
sp_invertebrate:09VT6	+	80.50	156.07	0.1564	1039	09VT6 drosophila melanogast
sp_plant:09SPF8	+	79.50	156.53	0.2144	715	09SPF8 arabidopsis thaliana (mc
sp_human:000509	+	78.00	162.18	0.3769	197	000509 homo sapiens (human). hy
sp_plant:09M815	+	76.00	148.69	0.5447	769	09M815 arabidopsis thaliana (mc
sp_invertebrate:09T616	-	74.00	153.69	1.10	200	09T616 locusta migratoria (m
sp_mammal:09T643	+	73.50	146.93	1.14	462	09T643 lama glama (llama). fib
sp_human:013344	+	72.00	142.87	1.67	528	013344 homo sapiens (human). fu
sp_bacteria:09L070	+	72.00	140.93	1.62	701	09L070 streptomyces coelicolor
sp_vertibrate:057480	+	72.00	138.06	1.53	1068	057480 gallus gallus (chicken)
sp_mammal:09T513	+	71.00	151.67	2.67	107	09T513 macaca mulatta (rhesus m
sp_mammal:09T655	+	71.00	151.23	2.65	114	09T655 macaca mulatta (rhesus m
sp_plant:09X1L1	+	71.00	145.68	2.40	257	09X1L1 arabidopsis thaliana (mc
sp_plant:09SU12	+	71.00	138.34	2.10	754	09SU12 arabidopsis thaliana (mc
sp_plant:004095	+	71.00	137.45	2.06	859	004095 arabidopsis thaliana (mc
sp_invertebrate:022739	+	70.50	147.42	2.88	171	022739 caenorhabditis elegans
sp_human:016273	+	69.00	146.36	4.49	126	016273 homo sapiens (human). fu
sp_human:09P266	+	69.00	143.37	4.25	195	09P266 homo sapiens (human). fu
sp_invertebrate:096500	+	68.50	137.94	4.49	371	096500 blaberus craniifer. ad
sp_plant:048709	+	68.50	133.55	4.14	705	048709 arabidopsis thaliana (mc
sp_human:09UJN4	+	68.00	144.26	5.87	129	09UJN4 homo sapiens (human). d
sp_plant:09X1L0	+	68.00	144.10	5.85	126	09X1L0 homo sapiens (human). hy
sp_invertebrate:09NB54	+	68.00	138.31	5.27	301	09NB54 arabidopsis thaliana (mc
sp_human:09UPN1	+	68.00	135.88	5.04	430	09NB54 dreisena polymorpha (
sp_vertibrate:013028	-	68.00	135.23	4.98	473	09UPN1 homo sapiens (human). ki
sp_bacteria:033828	-	67.50	133.71	5.65	507	013028 borreagius salda. antifi
sp_vertibrate:09PRH0	-	67.00	140.11	5.29	859	033828 thobacillus ferrooxidan
sp_plant:09SH19	-	67.00	144.38	8.00	91	09SH19 angulilla japonica (japan
sp_plant:09S159	+	67.00	146.83	6.97	275	09S159 trifolium aestivum (wheat
sp_plant:09SL9	+	66.50	134.90	7.85	313	09SL9 arabidopsis thaliana (mc
sp_invertebrate:09VX2	+	66.50	128.74	7.01	772	09SL9 arabidopsis thaliana (mc
sp_invertebrate:09UW2	+	66.50	125.59	6.62	1224	09UW2 drosophila melanogast
sp_invertebrate:022667	+	66.00	131.73	8.63	427	09UW2 aedes aegypti (yellow)
sp_invertebrate:024926	+	66.00	127.46	7.99	798	022667 caenorhabditis elegans
sp_vertibrate:024925	+	66.00	126.71	7.88	891	024926 eurytoma scolopes. per
sp_vertibrate:09PVF5	+	66.00	126.69	7.88	894	024925 eurytoma scolopes. per
sp_invertebrate:09NG18	+	65.00	124.64	7.59	1207	09PVF5 brachydanio rerio (zebr
		65.00	131.52	11.70	324	09NG18 drosophila similans (d

sp_invertebrate:09VE67	+	65.00	117.04	8.98	2703	09VE67 drosophila melanog
sp_invertebrate:061603	+	65.00	117.01	8.97	2715	061603 drosophila melanog
sp_invertebrate:077072	+	64.50	129.31	13.10	384	077072 eisenia foetida (co
sp_invertebrate:09Z201	+	64.50	129.29	13.09	385	09Z201 cavia porcellus (guin
sp_human:09P2N8	+	64.50	124.79	12.06	745	09P2N8 homo sapiens (human)

seq\_name: sp\_vertibrate:09PT52

#### seq\_documentation\_block:

ID	Q9PT52	PRELIMINARY:	PRT:	263 AA.
AC	Q9PT52:			
DT	01-MAY-2000 (TREMUREL_13, Created)			
DT	01-MAY-2000 (TREMUREL_13, Last sequence update)			
DT	01-JUN-2000 (TREMUREL_14, Last annotation update)			
DE	BPP-CNP PRECURSOR HOMOLOG.			
OS	Agkistrodon blomhoffi (Mamushi) (Gloydus blomhoffii).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;			
OC	Viperidae; Crotalinae; Agkistrodon.			
OX	NCBI_TaxID=61300;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=VENOM GLAND;			
RA	Murayama N.;			
RT	"Agkistrodon blomhoffi cDNA for BPP-CNP precursor homolog."			
RL	Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AB020810; BAA36953.1;			
DR	INTERPRO; IPR000663;			
DR	PFAM; PF00212; ANP.1.			
DR	PRINTS; PR00710; NATPEPTIDES.			
DR	PROSITE; PS00263; NATRURETIC_PEPTIDE.1.			
FT	CHAIN	31	41	BLOMOTIN.
FT	CHAIN	31	40	POTENTIATOR A.
FT	CHAIN	49	59	LEU3-BLOMOTIN.
FT	CHAIN	67	77	POTENTIATOR C.
FT	CHAIN	85	95	POTENTIATOR B.
FT	CHAIN	103	113	POTENTIATOR B.
FT	CHAIN	117	127	POTENTIATOR E.
FT	CHAIN	242	263	CNP-22.
SO	SEQUENCE	263 AA;	27339 MW;	407BA9A572BF5FC8 CRC64;

#### alignment\_scores:

Quality:	95.50	Length:	63
Ratio:	3.183	Gaps:	3
Percent Similarity:	47.619	Percent Identity:	41.270

#### alignment\_block:

US-09-306-689-10 x Q9PT52 ..

Align seg 1/1 to: Q9PT52 from: 1 to: 263

1	CAGCATTTGGAGCTACGGCTCGCCCTGACAGCGT.....	36
27	GINGINTPSERGNGIYArgProProGlyProProIleProArgLeuVa	43
37	TCTCAAGATTGAGCTACGGCTCGCCCTGAGTCT.....	75
43	IValGININTPSERGNGIYLeuProProGlyProProIleProArgL	60
76	.....AGCCAGATTGAGCTACGGCTCGCCCTGACAGCGT.....	114
60	eUValValGININTPSERGNGIYLeuProProGlyProProIlePro	76
115	.....AGCCAGATTGAGCTACGGCTCGCCCTGACAGCGT.....	144
77	ProLeuValValGININTPSERGNGIYLeuProPro	89

seq\_name: sp\_invertebrate:09WTY8  
seq\_documentation\_block:  
ID Q9WTY8 PRELIMINARY; PRT; 836 AA.  
AC Q9WTY8;



```
alignment_scores:
  quality: 78.00
  length: 83
```

Ratio: 2.438 Gaps: 3  
Percent Similarity: 38.554 Percent Identity: 27.711

alignment\_block:

US-09-306-689-10 x 000509 ..

Align seg 1/1 to: 000509 from: 1 to: 197

```

1 CAGCATTTGAGCTACGGCTGCGCTGGCAGCGGT..... 36
   ||| ||||| ..... |||||
61 GlnProtrpSerSerSerSerProGlyProArghisSerAlaLeuAs 77
   :||| ||||| ..... |||||
37 .....TCTCAGATTGAGCTACGGCTGCGCTGGCAGCGGTCT... 75
   :||| ||||| ..... |||||
77 pLeuValThrGlnProtrpSerSerSerSerProGlyProArghis 94
   :||| ||||| ..... |||||
76 .....AGCCAGCATTTGAGCTACGGCTGCGCTGGCAGCGGT 108
   :||| ||||| ..... |||||
94 eAlaLeuValLeuIleThrGlnProtrpThrSerSerLeuSerProGly 110
   :||| ||||| ..... |||||
109 AGCGGT..... 114
111 ProProHisSerAlaLeuAspLeuIleThrHisProGlyProHisLe 127
115 .....AGCCAGATTGAGCTACGGCTGCGCTGGCAGCGGT 147
   :||| ||||| ..... |||||
127 uAlaLeuAspLeuIleThrGlnProtrpSerTrpSerLeuSerProGly 143

```

seq\_name: sp\_plant:09M815

seq\_documentation\_block:

```

ID 09M815 PRELIMINARY; PRT; 769 AA.
AC 09M815;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE PUTATIVE PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE.
GN TPPI.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Rouning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Niernman W.C., Fraser C.M.;
RT *Arabidopsis thaliana chromosome I BAC TPPI genomic sequence.*;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC018908; AAF27664.1; -.
KW kinase.
SQ SEQUENCE 769 AA; 87585 MW; 652P44182F493845 CRC64;

```

alignment\_scores: Quality: 76.00 Length: 77  
Ratio: 2.714 Gaps: 2  
Percent Similarity: 36.364 Percent Identity: 29.870

alignment\_block:

US-09-306-689-10 x 09M815 ..

Align seg 1/1 to: 09M815 from: 1 to: 769

```

7 TGAAGCTACGGCTGCGCTGGCAGCGGTCTCAAGATTGAGCTACGG 56
   ||| ||||| ..... |||||
116 TrpArgGlnGlyLeuGlnAspGlySerGlySerTrpTyrAsnG1 132
   :||| ||||| ..... |||||
57 CCTGCGTCCGGGTGCTCT..... 75
132 yAsnArgPheIleGlyAsnTrpLysGlyLysMetSerGlyArgGly 149

```

```

76 .....AGCCAGCATTTGAGCTACGGC 96
   :||| ||||| ..... |||||
149 alMetSerTrpAlaAsnGlyAspLeuPheAsnGlyPheTrpLeuAsnGly 165
   :||| ||||| ..... |||||
97 CTGGCGCTGGCAGCGGTAGCCAA..... 120
   :||| ||||| ..... |||||
166 LeuArgHisGlySerGlyValTyrLysTyrAlaAspGlyGlyPheTyrPh 182
   :||| ||||| ..... |||||
121 .....GATTGAGCTACGGCTGCGCTGGCAGCGGT 147
   :||| ||||| ..... |||||
182 eGlyThrTrpSerArgGlyLeuLysAspGly 192

```

seq\_name: sp\_invertebrate:p82166

seq\_documentation\_block:

```

ID P82166 PRELIMINARY; PRT; 200 AA.
AC P82166;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CUTICLE PROTEIN 19.8 (LMNCP19.8).
OS Locusta migratoria (migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=FIFTH INSTAR LARVAE CUTICLE;
RA Anderson S.O.;
RL Submitted (NOV-1999) to the SWISS-PROT data bank.
CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF MIGRATORY LOCUST WHICH
CC -1- CONTAINS MORE THAN 100 DIFFERENT STRUCTURAL PROTEINS
CC -1- DOMAIN: THE TETRAPEPTIDE (A-A-P-[AV]) REPEATS FOUND THROUGHOUT THE
CC PROTEIN ARE ALSO PRESENT IN MANY PROTEINS CONSTITUTING THE
CC PROTECTIVE ENVELOPE OF OTHER SPECIES.
CC -1- SIMILARITY: CONTAINS A CUTICLE CONSENSUS DOMAIN.
DR INTERPRO: IPR000618; -.
DR PIRAM: PR00379; Insect_cuticle; 1.
DR PRINTS: PR00947; CUTICLE.
DR PROSITE: PS00233; CUTICLE; 1.
KW Structural protein; Cuticle; Repeat.
FT DOMAIN 20 180 8 X 4 REPEATS.
FT REPEAT 20 23 1.
FT REPEAT 25 29 2.
FT REPEAT 35 43 3.
FT REPEAT 46 54 4.
FT REPEAT 57 65 5.
FT REPEAT 76 84 6.
FT REPEAT 95 103 7.
FT REPEAT 114 122 8.
FT REPEAT 133 141 8.
FT REPEAT 152 160 8.
FT REPEAT 171 179 8.
SQ SEQUENCE 200 AA; 19815 MW; 3802F2DB3BA6F92D CRC64;

```

alignment\_scores: Quality: 74.00 Length: 47  
Ratio: 2.387 Gaps: 2  
Percent Similarity: 65.957 Percent Identity: 46.809

alignment\_block:

US-09-306-689-10/rev x P82166 ..

Align seg 1/1 to: P82166 from: 1 to: 200

```

144 CGAAGCAGAGCGCTAGCTCAATTTGGCTACCGCT.....GCCAGGC 101
   ||||| ..... |||||
140 ArgThrValAlaAlaProAlaValAlaAlaProAlaValArgAl 156
   :||| ||||| ..... |||||
100 GCAGCGCTAGCTCCATGCTGGGTAGAGCCAGCGAGCGAGCGCTAG 51
   :||| ||||| ..... |||||
156 aAlaIleAlaAlaProAlaTyrAla.....ThryAlaAla 169

```





58 InlathrtYrGlyLeu 63

seq\_name: sp\_mammal:097655

seq\_documentation\_block:  
ID 097655 PRELIMINARY; PRT; 114 AA.

DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)  
DE PRECURSOR.  
GN GNRH2.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]

RP SEQUENCE FROM N.A.

RA White R.B., Urbanski H.F., Fernald R.D.;

RT "A second gene for gonadotropin-releasing hormone is expressed in the  
rhesus macaque (Abstract #632.18)."

RL Abstr. - Soc. Neurosci. 24:1609-1609(1998).

CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
FOLLICLE-STIMULATING HORMONES.  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
DR EMBL; AF097356; AAD09106.1; -.  
DR INTERPRO; IPR002012; -.  
DR PFAM; PF00446; GNRH; 1.  
DR PROSITE; PS00473; GNRH; 1.  
DR PRODOM; PD029787; -; 1.  
KW Hormone; Amidation.  
SQ SEQUENCE 114 AA; 12533 MW; 8B70D690D5BD5103 CRC64;

alignment\_scores:

Quality:	71.00	Length:	39
Ratio:	2.731	Gaps:	2
Percent Similarity:	66.667	Percent Identity:	46.154

alignment\_block:  
US-09-306-689-10 x 097655 ..

Align seg 1/1 to: 097655 from: 1 to: 114

```

1 CAGCATTTGAGCTACGGCTGCGCCCTGCGC.....AGCGG 35
|||||
25 GlnHISTPserHisGlyTrpTyrProGlyGlySArgAlaLeuSer 41
|||||
36 TTCTCAAGATTGAGCTACGGCTGCGT...CCGGGTGGCTCTAGCCAGC 82
|||||
41 rAlaGlnAspProGlnAsnAlaLeuArgProAlaGlySerProAlaG 58
|||||
83 ATTGAGCTACGGCCTG 99
|||||
58 InlathrtYrGlyLeu 63

```

seq\_name: sp\_plant:09XII1

seq\_documentation\_block:

ID 09XII1 PRELIMINARY; PRT; 257 AA.

AC 09XII1;

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE F19G14.22 PROTEIN.

GN F19G14.22.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV, COLUMBIA.

RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,

RA Barnstead M.E., Mason T.M., Bowman C.L., Renning C.M., Benito M.,

RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,

RA Fraser C.M., Venter J.C.;

RT "Arabidopsis thaliana chromosome II BAC F19G14 genomic sequence.";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC006438; AAD41986.1; -.

DR HSSP; P00303; 2CBP.

SQ SEQUENCE 257 AA; 27400 MW; DA59B7A93D53A17 CRC64;

alignment\_scores:

Quality:	71.00	Length:	69
Ratio:	2.152	Gaps:	3
Percent Similarity:	47.826	Percent Identity:	26.087

alignment\_block:  
US-09-306-689-10 x 09XII1 ..

Align seg 1/1 to: 09XII1 from: 1 to: 257

```

7 TGGAGCTAGCGCTGCGCCCT.....GGCAGCGGCTTC 38
|||||
87 TTPGlyTTPGlyGlyValProAsnAsnThrHisSerSerglySergly 103
|||||
39 TCAAGATTG.....AGCTACGGCTGCGCGGGTG 70
|||||
103 ySerglyTTPGlyMetGlyProAsnAsnAsnTyrSerglySergly 120
|||||
71 GCTTACGACGAGCATTGAGCTACGGCTGCGCCCTGCGCAGC..... 111
|||||
120 ySerglySerglyTTPGlyTyrGlyHisSerTyrAsnAla 136
|||||
112 .....GTAGCAAGATTGAG 128
|||||
137 ThrTyrAsnGlyProArgLysIleIleValGlyGlyAspLysGlyTTPH 153
|||||
129 CTACGAGC 135
|||||
153 TTYrGly 155

```

**THIS PAGE BLANK (uspto)**







```

FT PEPTIDE 6 15 GONADOLIBERIN I
FT ACT_SITE 19 >67 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT MOD_RES 8 8 ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 15 15 SIMILARITY).
FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
FT NON_TER 67 67 SIMILARITY).
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

alignment_scores:
Quality: 76.50 Length: 19
Ratio: 4.500 Gaps: 1
Percent Similarity: 89.474 Percent Identity: 78.947

alignment_block:
US-09-306-689-10 x GONL_MACMU ..

Align seg 1/1 to: GONL_MACMU from: 1 to: 67

.67 GGAGGCTGAGCCAGCATGGAGTACGCGCTGCGCCGAGCAGGGTAG 116
11 |||||
2 GlycSerSerGlnHisTrpPertyrGlyLeuAlaArgProGly..GlyLy 17
117 CCAGAT 123
:::|
17 sarGasp 19

Seq_name: Swissprot_39:GONL_HUMAN

seq_documentation_block:
ID GONL_HUMAN STANDARD; PRT; 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED
DE PEPTIDE I].
GN GNRH1 OR GNRH OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=8936682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RL hormone gene.";
RN Nucleic Acids Res. 17:6403-6403(1989).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
of gonadotropin-releasing hormone and prolactin release-inhibiting
factor in human and rat.";
RN Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RL releasing hormone.";
RN Nature 311:666-668(1984).
[4]
RN SEQUENCE OF 24-33.
RX MEDLINE=83126573; PubMed=6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide

```

```

RT biosynthesized in the human placenta."
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACIREL (AYERST LABS),
CC LUTREPULE OR LUTRELEF (FERRING PHARMACEUTICALS) AND RELISORM
CC (SERONO).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X01059; CAA25526.1; -.
DR EMBL, M12578; AAA35916.1; -.
DR EMBL, X15215; CAA33285.1; -.
DR PIR, A01410; RHUG.
DR PIR, A26173; A26173.
DR PIR, S05308; S05308.
DR MIM, 152760; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GNRH, 1.
DR PROSITE: PS00473; GNRH, 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal.
KW
FT SIGNAL 1 23
FT CHAIN 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT CONFLICT 16 16 W -> S (IN REF. 3).
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

alignment_scores:
Quality: 76.50 Length: 19
Ratio: 4.500 Gaps: 1
Percent Similarity: 89.474 Percent Identity: 78.947

alignment_block:
US-09-306-689-10 x GONL_HUMAN ..
Align seg 1/1 to: GONL_HUMAN from: 1 to: 92

67 GGTGGCTTACCCAGCATTTGGAGCTACGGCCCTGGCCCTGGACGGCGTAG 116
111 |||||||
20 GlyCysSerSerClnHisTrpSerTyrglyLeuArgProGly... Glyly 35
117 CCAAGAT 123
35 sarGasp 37

seq_name: SwissProt_39:GONL_MOUSE
seq_documentation_block:
ID GONL_MOUSE STANDARD; PRT; 90 AA.
AC P13562.
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
DE I].
DE GNRH1 OR GNRH.

```



AC P07490;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
 DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR  
 DE I].  
 GN GNRH1 OR GNRH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-86094338; PubMed-2867548;  
 RA Adelman J.P., Mason A.J., Haylick J.S., Seeburg P.H.;  
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor  
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
 RT factor in human and rat.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-89384661; PubMed-2476669;  
 RA Bond C.T., Haylick J.S., Seeburg P.H., Adelman J.P.;  
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and  
 RT hypothalamic expression.";  
 RL Mol. Endocrinol. 3:1257-1262(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-THYMUS;  
 RX MEDLINE-93105480; PubMed-1468115;  
 RA Maher C.G., Marchetti B., Lebouef R.D., Blalock J.E.;  
 RT "Thymocytes express a mRNA that is identical to hypothalamic  
 RT luteinizing hormone-releasing hormone mRNA.";  
 RL Cell. Mol. Neurobiol. 12:447-454(1992).  
 RN [4]  
 RP SEQUENCE OF 1-47 FROM N.A.  
 RC TISSUE-HEART;  
 RX MEDLINE-87149087; PubMed-3547652;  
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;  
 RT "Two mammalian genes transcribed from opposite strands of the same  
 RT DNA locus.";  
 RL Science 235:1514-1517(1987).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; S50870; AAB24572.1; -  
 DR EMBL; M12579; AAA41263.1; -  
 DR EMBL; M31670; AAA41264.1; -  
 DR EMBL; M15527; AAA42141.1; ALT\_SEQ.  
 DR EMBL; M15528; AAA42139.1; -  
 DR EMBL; M15528; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; B26173; RHRTG.  
 DR PIR; A48410; A48410.  
 DR INTERPRO; IPR002012; -  
 DR PFM; PF00446; GNRH.1.  
 DR PROSITE; PS00473; GNRH.1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placental; Signal.  
 FT SIGNAL 1 23  
 FT CHAIN 24 92 PROGONADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.

FT PEPTIDE 37 92 PROLACTIN RELEASE-INHIBITING FACTOR I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT MOD\_RES 24 24 ACTIVITY.  
 FT MOD\_RES 33 33 PYROLIDONE CARBOXYLIC ACID.  
 SO SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;  
 alignment\_scores:  
 Quality: 74.00 Length: 14  
 Ratio: 5.692 Gaps: 0  
 Percent Similarity: 92.857 Percent Identity: 92.857  
 alignment\_block:  
 US-09-306-689-10 x GON1\_RAT ..  
 Align seg 1/1 to: GON1\_RAT from: 1 to: 92  
 67 GGTGGCTCTAGCCAGCATGAGCTACGGCCCTGCGCCCTGGC 108  
 ||| ||||||||| ||||||||| ||||||||| |||||||||  
 20 GlyCysSerSerGlnHisIstperItyrGlyLeuArgProGly 33  
 seq\_name: SwissProt\_39:SVS2\_RAT  
 seq\_documentation\_block:  
 ID SVS2\_RAT STANDARD; PRT; 414 AA.  
 AC P22006;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE SEMINAL VESICLE SECRETORY PROTEIN II PRECURSOR (SVS II).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE-90277684; PubMed-2351680;  
 RA Harris S.E., Harris M.A., Johnson C.M., Bean M.F., Dodd J.G.,  
 RA Marcus R.J., Carr S.A., Crabb J.W.;  
 RT "Structural characterization of the rat seminal vesicle secretion II  
 RT protein and gene.";  
 RL J. Biol. Chem. 265:9896-9903(1990).  
 CC -1- FUNCTION: THE RAT SEMINAL VESICLE CONTAINS SIX MAJOR ANDROGEN-  
 CC DEPENDENT SECRETORY PROTEINS REFERRED TO AS SVS I-VI. THE SVS  
 CC I-III PROTEINS APPEAR TO BE COMPONENTS OF THE RAT COPULATORY  
 CC PLUG, WITH THE SVS II PROTEIN BEING THE MAJOR COMPONENT.  
 CC -1- PTM: THE REPEATING UNIT APPEARS TO BE INVOLVED IN THE FORMATION OF  
 CC THE COPULATORY PLUG VIA A TRANSGLYUTAMINASE REACTION CROSS-LINKING  
 CC GLUTAMINE AND LYSINE RESIDUES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; J05443; AAA42192.1; -  
 DR PIR; A36443; A36443.  
 DR INTERPRO; IPR002080; -  
 DR PROSITE; PS00515; SVP\_II; 12.  
 KW Testosterone; Seminal vesicle; Signal; Repeat; Copulatory plug.  
 FT SIGNAL 1 22  
 FT CHAIN 23 414  
 FT MOD\_RES 23 23 SEMINAL VESICLE SECRETORY PROTEIN II.  
 FT MOD\_RES 108 311 PYROLIDONE CARBOXYLIC ACID.  
 FT REPEAT 108 120 13 X 13 AA TANDEM REPEATS.  
 FT REPEAT 127 139 1.  
 FT REPEAT 140 152 2.  
 FT REPEAT 153 165 3.  
 FT REPEAT 166 178 4.  
 FT REPEAT 179 191 5.  
 FT REPEAT 179 191 6.











```

RT Picus.";
RL Dev. Biol. 148:473-480(1991).
CC -I- FUNCTION: MAJOR MATRIX PROTEIN OF THE SEA URCHIN EMBRYO SPICULE
CC WHICH DIRECTS CRYSTAL GROWTH IN CERTAIN ORIENTATIONS AND INHIBIT
CC GROWTH IN OTHERS.
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- TISSUE SPECIFICITY: EMBRYO SPICULE.
CC -I- DOMAIN: THE REPETITIVE DOMAIN MAY PROVIDE A CALCITE BINDING
CC MATRIX.
CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -I- SIMILARITY: BELONGS TO THE SM50 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL; X59616; CAA42179.1; -.
DR INTERPRO; IPR001304; -.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Matrix protein; Signal; Repeat.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 335 34 KDA SPICULE MATRIX PROTEIN.
FT DOMAIN ? ? C-TYPE LECTIN.
FT DOMAIN 131 178 PRO-RICH.
SO SEQUENCE 335 AA; 35426 MW; AA032281DDB0AE97 CRC64;

```

alignment_scores:		
Quality:	68.00	Length: 46
Ratio:	2.000	Gaps: 2
Percent Similarity:	73.913	Percent Identity: 39.130

```
alignment_block:
US-09-306-689-10 x SM34_LYTP1 .
```

Align seg 1/1 to: SM34\_LYTP1 from: 1 to: 335

[illegible]

seq\_name: SwissProt\_39:GON1\_XENLA

seq\_documentation\_block:

ID	GONI_XENLA	STANDARD;	PRT;	89 AA.
----	------------	-----------	------	--------

DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GONADOLIBERIN 1 PRECURSOR (GONADOTROPIN-RELEASING HORMONE 1) (GNRH-1)  
DE (LH-RH) (LUTALIBERIN 1).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
OC Xenopodinae; Xenopus.

RP SEQUENCE FROM N.A.  
RC TISSUE=FOREBRAIN;  
RX MEDLINE=94185563; PubMed=8137750;  
RA Hayes W.P., Wray S., Battey J.F.;

RT		"The gonadotropin-releasing hormone-I (GNRH-I) gene has a mammalian-like expression pattern and conserved domains in GNRH-associated peptide, but brain onset is delayed until metamorphosis."
RL		Endocrinology 134:1835-1844(1994).
CC	-I FUNCTION:	STIMULATES THE SECRETION OF GONADOTROPINS.
CC	-I SIMILARITY:	BELONGS TO THE GNRH FAMILY.
CC		-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC		-----
DR	EMBL; L28040;	AAA49728.1; -.
DR	INTERPRO;	IIPRO02012; -.
DR	PFBM;	PFO0446; GNRH; 1.
DR	PROSITE;	PS00473; GNRH; 1.
KW	Cleavage	on pair of basic residues; Hormone; Amidation; Hypothalamus; Signal.
KW	SIGNAL	
FT	CHAIN	1 23 PROGONADOLIBERIN I.
FT	CHAIN	24 89 GONADOLIBERIN I.
FT	PEPTIDE	24 33 CONADOTROPIN I.
FT	CHAIN	37 89 GONADOTROPIN-RELEASING HORMONE ASSOCIATED PEPTIDE.
FT	PEPTIDE	37 85 GNRH-ASSOCIATED PEPTIDE I (GAP).
FT	MOD_RES	24 24 PYRROLIDONE CARBOXYLIC ACID.
FT	MOD_RES	33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SO	SEQUENCE	89 AA; 10246 MW; 6FAFF36FBAED04284 CRC64;

alignment_scores:		
Quality:	66.50	Length: 16
Ratio:	4.433	Gaps: 1
Percent Similarity:	93.750	Percent Identity: 75.000

```
alignment_block:
```

Align seg 1/1 to: GON1\_XENLA from: 1 to: 89

```

76 AGCCAGCATTTGGACCTACGGGCTTCGGCCCTGGGACAGGGTACGCCAAGAT 12
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
23 ALAAGNHSTRSPERTYRGLYLEAHARGPROGLY...GLYLysARGasp 37
seq_name: SwissProt_39::ISOA_PSEAY

```

seq\_name: SwissProt\_39:ISOA\_PSEAY

seq\_documentation\_block:

ID	ISOA_PSEAY	STANDARD;	PRT;	776 AA
1	110340			

AC P10342;  
DT 01-MAY-71

DT 01-MAR-1  
DT 15-DEC-1

DT 15-DEC-1

DE ISOAMYLA

GN I AM.

OS Pseudomo

OC Bacteria

RN [1]  
PD  
CONFERENCE

BC STRAIN

AC CINCINNATI  
RX MEDLINE=

RA Amemura

## RT "Cloning"

RT Pseudomo

J. Biol.

RN [2]  
PD EFFORTS

RP	SEQUENCE
BC	STRATN=T

RX MEDLINE=

RA Chen J. H.

RT	<sup>14</sup> Nucleot
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

RT isoamylase-hyperproducing mutant, Pseudomonas amyloclavata JD210.  
RL Biochim. Biophys. Acta 1087:309-315(1990).  
[3]  
RC SEQUENCE OF 744-776 FROM N.A.  
RX STRAIN-SB-15;  
RA MEDLINE-89327147; PubMed-2753857;  
RA Amenura A., Fujita M., Futai M.;  
RT "transcription of the isoamylase gene (lam) in Pseudomonas  
RT amyloclavata SB-15.";  
RL J. Bacteriol. 171:4320-4325(1989).  
[4]  
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE-98387895; PubMed-9719642;  
RA Katsuya Y., Mezaki Y., Kubota M., Matsuura Y.;  
RT "Three-dimensional structure of Pseudomonas isoamylase at 2.2-A  
RT resolution.";  
RL J. Mol. Biol. 281:885-897(1998).  
CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,6-ALPHA-D-GLUCOSIDIC BRANCH  
CC LINKAGES IN GLYCOGEN, AMYLOPECTIN AND THEIR BETA-LIMITS DEXTRINS.  
CC -I- INDUCTION: BY MALTOSYL.  
CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: J03871; AAA25854.1;  
DR EMBL: X13378; CAA31754.1;  
DR PIR: A28109; A28109.  
DR PDB: 1BR2; 12-AUG-98.  
DR INTERPRO: IPR000461;  
DR PFAM: PF00128; alpha-amylose; 1.  
DR KW Hydrolyase; Glycosidase; Signal; 3D-structure.  
FT SIGNAL 1 26  
FT CHAIN 27 776  
FT ACT\_SITE 401 401  
FT ACT\_SITE 481 481  
FT ACT\_SITE 536 536  
FT DISULFID 410 422  
FT DISULFID 546 616  
FT DISULFID 738 766  
FT CONFLICT 8 8  
FT CONFLICT 126 126  
FT CONFLICT 169 171  
FT CONFLICT 386 386  
FT CONFLICT 413 416  
FT CONFLICT 454 489  
FT CONFLICT 555 556  
FT CONFLICT 650 657  
SQ SEQUENCE 776 AA: 83626 MW: F738BF8040246169 CRC64;  
A -> G (IN REF. 1).  
F -> C (IN REF. 1).  
GAS -> AH (IN REF. 1).  
L -> V (IN REF. 1).  
CAVT -> AVH (IN REF. 1).  
SGIDLEAEPAIGNSYLGFGPGGSEWNGLFKDS ->  
TWICICRNIGPSAATRTSWVDSRRVRYVEMSVPRQ (IN  
REF. 1).  
WF -> S (IN REF. 1).  
AFKRAHPA -> RSARHIP (IN REF. 1).

alignment\_scores: Quality: 66.00 Length: 37  
Ratio: 2.750 Gaps: 1  
Percent Similarity: 64.865 Percent Identity: 35.135

alignment\_block:

US-09-306-689-10 x ISOA\_PSEAY ..

Align seg 1/1 to: ISOA\_PSEAY from: 1 to: 776

13 TACGGCTGGCGCTGAGGCTTCACAGATGAGAGCTACGGCTGCG 62  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
544 TysercysannglyalaasnasnserglinalatprrprrglyProse 560

63 TCCGGGTGGCTTAGCCAGCAT.....TGGAGCTACGGCTGCGCCCTG 106  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
560 RASpGLyGLyThrsrThrsnTySerTrpAspGLnGLyMetSerAlag 577  
107 GCACGGGTAGC 117  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
577 LyThrGLyAla 580



```

A:Status: preliminary: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-67 <RES>
A:Cross-references: GB:S75918; NID:g912831; PIDN:AA833096.1; PID:g912832
C:Superfamily: gonadolibrin

alignment_scores:
      Quality: 76.50      Length: 19
      Ratio: 4.500      Gaps: 1
      Percent Similarity: 89.474      Percent Identity: 78.947

alignment_block:
  US-09-306-689-10 x I78541 ..

  Align seg 1/1 to: I78541 from: 1 to: 67

      67 GGTGGCTTACCCAGCATTTGAGCTACGGCTTGCCCTCGCGACCGGTAG 116
          ||| ||||||||||||||||||||||||||||| |||::
      2 GlycylSerSerGlnHisTrpSerTyrGlyLeuArgProGly... Glyly 17

      117 CCAGAT 123
          |||::|||
      17 SATGASP 19

seq_name: p1rl:RHHUG

seq_documentation_block:
  gonadolibrin precursor - human
  N:Alternate names: gonadotropin releasing hormone (GNRH); luteinizing hormone releasing
  N:Comments: gonadolibrin-associated protein (GAP); progonadolibrin
  C:Species: Homo sapiens (man)
  C:Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
  C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
  R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
  Nucleic Acids Res. 17: 6403-6404, 1989
  A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone ge
  A:Reference number: S05308; MUID:89366682
  A:Status: translation not shown
  A:Accession: S05308
  A:Molecule type: DNA
  A:Residues: 1-92 <NAV>
  A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
  R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
  Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
  A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
  A:Reference number: A94090; MUID:86094338
  A:Accession: A26173
  A:Molecule type: mRNA
  A:Residues: 1-92 <ADP>
  A:Cross-references: GB:M12578; NID:g183418; PIDN:AA35916.1; PID:g386749
  A:Experimental source: hypothalamus
  R:Seeburg, P.H.; Adelman, J.P.
  Nature 311, 666-668, 1984
  A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing h
  A:Reference number: A93342; MUID:85012739
  A:Accession: A93342
  A:Molecule type: mRNA
  A:Residues: 1-15, 'S', 17-92 <SEE>
  A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
  A:Experimental source: Placenta
  R:Tan, L.; Roussseau, P.
  Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
  A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
  A:Reference number: A90108; MUID:83126573
  A:Accession: A90108
  A:Molecule type: protein
  A:Residues: 24-33 <TAN>
  A:Experimental source: Placental trophoblasts
  R:Leibovitz, D.; Koch, T.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterd
  FBS Lett. 346, 203-206, 1994
  A:Title: Sequential degradation of the neurotrophin-releasing hormone by th
  A:Reference number: S45718; MUID:94283597

```

```

A:Contents: annotation; degradation pathway of synthetic hormone
C:Genetics:
A:Gene: GDB:GNRH, LHRH, GRH
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A:Map position: bp21-8p11.2
A:Introns: 47/3; 79/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitrop
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-33/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progonaoliberin #status predicted <PGN>
F:23-92/Product: gonadoliberin #status experimental <MAT>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

alignment_scores:
Quality: 76.50 Length: 19
Ratio: 4.500 Gaps: 1
Percent Similarity: 89.474 Percent Identity: 78.947

alignment_block:
US-09-306-689-10 x RHNG ..
Align seg 1/1 to: RHNG from: 1 to: 92

        67 GGTCCTAGCCACATTCAGACTACGGCTCGGCCCTGCAGCGTAG 116
           ||| | | | | | | | | | | | | | | | | | | | | |
          20 GLYCYSERSErglnHisTrpSerTyrGlyLeuAcgProGly...GlyLy 35
              117 CCAGAT 123
                ::::|||
               35 sArgAsp 37

seq_name: pLr:RHMSG

seq_documentation_block:
gonadoliberin precursor - mouse
N:Alternate names: gonadotropin-releasing hormone (GNRH); luteinizing hormone release
N:Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: A47578
R:Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli
Science 234, 1366-1371, 1986
A>Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible
A:Reference number: A47578; MUID:87069928
A:Accession: A47578
A:Molecule type: DNA
A:Residues: 1-90 <MAS>
A:Cross-references: EMBL:M4872; MID:g193576; PIDN:AAA37717.1; PID:g387175
C:Genetics:
A:Introns: 45/3; 77/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitrop
A>Note: gonadoliberin-associated protein may have prolactin release inhibiting activi
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-33/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin #status predicted <GLB>
F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

alignment_scores:
Quality: 74.00 Length: 14
Ratio: 5.692 Gaps: 0
Percent Similarity: 92.857 Percent Identity: 92.857

```

alignment\_block:  
US-09-306-689-10 x RHMSG

Align seg 1/1 to: RHMSG from: 1 to: 90

67 GGTCGCTTACGACGATTCGAGCTACGGCTGCGCCCTGCG 108  
|||||  
18 GlycylserSerGlnHisTrpSerTyrGlyLeuArgProGly 31

seq\_name: p1r1:RHRTG

seq\_documentation\_block:

gonadoliberin precursor - rat

N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormo

N:Contains: gonadoliberin; prolactin release-inhibiting factor

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1999

C:Accession: A40147; B26173; A48410

R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.

Mol. Endocrinol. 3, 1257-1262, 1989

A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex

A:Reference number: A40147; MUID:89384661

A:Accession: A40147

A:Residues: 1-92 <BON>

A:Cross-references: GB:M1670; NID:g204447; PIDN:AAA1264.1; PID:g204448

R:Adelman, J.P.; Mason, A.U.; Hayflick, J.S.; Seeburg, P.H.

Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986

A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot

A:Reference number: A94090; MUID:86094338

A:Accession: B26173

A:Molecule type: mRNA

A:Residues: 1-92 <ADP>

A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA1263.1; PID:g204446

R:Meier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.

Cell. Mol. Neurobiol. 12, 447-454, 1992

A:Title: Tumor cells express a mRNA that is identical to hypothalamic luteinizing hormone

A:Reference number: A48410; MUID:93105480

A:Accession: A48410

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-92 <MAI>

A:Cross-references: GB:S50870; NID:g262059; PIDN:AAH24572.1; PID:g262060

A:Experimental source: thymus

A>Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIIP:121083)

C:Genetics:

A:Introns: 47/3; 79/3

C:Function:

A:Description: stimulates pituitary secretion of luteotropin and follitropin

C:Superfamily: gonadoliberin-associated protein may have prolactin release inhibiting activity

C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid; x

F:1-33/Domains: signal sequence #status predicted <SIG>

F:24-92/Product: progadoliberin #status predicted <P>

F:24-33/Product: gonadoliberin #status predicted <GIN>

F:37-92/Product: prolactin release-inhibiting factor #status predicted <PRI>

F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted

F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

alignment\_scores:

Quality:	74.00	Length:	14
Ratio:	5.692	Gaps:	0
Percent Similarity:	92.857	Percent Identity:	92.857

alignment\_block:

US-09-306-689-10 x RHRTG

Align seg 1/1 to: RHRTG from: 1 to: 92

67 GGTCGCTTACGACGATTCGAGCTACGGCTGCGCCCTGCG 108  
|||||  
20 GlycylserSerGlnHisTrpSerTyrGlyLeuArgProGly 33

seq\_name: p1r2:G02127

seq\_documentation\_block:

tus-like protein - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 23-Jul-1999

C:Accession: G02127

R:Itou, K.; Kawase, M.

submitted to the EMBL Data Library, September 1995

A:Reference number: G09199

A:Accession: G02127

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-528 <ITD>

A:Cross-references: EMBL:U35561; NID:g1040969; PIDN:AAA79948.1; PID:g1040970

C:Superfamily: RNA-binding protein EWS; ribonucleoprotein repeat homology

F:289-364/Domains: ribonucleoprotein repeat homology <RNM>

alignment\_scores:

Quality:	72.00	Length:	50
Ratio:	2.769 <td>Gaps:</td> <td>2</td>	Gaps:	2
Percent Similarity:	52.000 <td>Percent Identity:</td> <td>40.000</td>	Percent Identity:	40.000

alignment\_block:

US-09-306-689-10 x G02127

Align seg 1/1 to: G02127 from: 1 to: 528

13 TACGGCTGCGGCT.....GGCAGCG 35  
||||| :|||  
99 TyrglyProGlnProThrProSerSerThrSerGlySerTyrIleSer 115  
:|||||  
36 TTCTCAAGATTGAGCTACGCGCTGCGGCTGAGCTACCAAGATT 85  
:||||| :|||  
115 SerGlnThrSerSerTyrGlyGlnPro..... 125  
:|||||  
86 GGAGCTACGCGCTGCGGCTGCGAGCGGTTAGCCAGATTGAGCTACGCG 135  
:||||| :|||  
126 ..SerTyrGlyGlnProSerTyrGlyGlyGlnGlnSerTyrGly 141  
:|||||

seq\_name: p1r2:T42382

seq\_documentation\_block:

guanylate cyclase (EC 4.6.1.2) 1, retinal - chicken (fragment)

N:Alternate names: guanylyl cyclase; retinal guanylate cyclase 1

C:Species: Gallus gallus (chicken)

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jun-2000

C:Accession: T42382

R:Semple-Rowland, S.L.; Lee, N.R.; Van Hooser, J.P.; Palczewski, K.; Baehr, W.

Proc. Natl. Acad. Sci. U.S.A. 95, 1271-1276, 1998

A:Title: A null mutation in the photoreceptor guanylate cyclase gene causes the retin

A:Reference number: Z22141; MUID:98115910

A:Accession: T42382

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1068 <SEM>

A:Cross-references: EMBL:AF036942; NID:g2828021; PID:g2828022; PIDN:AAC24560.1

A:Experimental source: strain Rhode Island Red; retina-pigment epithelium-choroid

C:Genetics:

A:Gene: GC1

C:Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain h

C:Keywords: cGMP biosynthesis; glycoprotein; phosphorus-oxygen lyase; transmembrane p

alignment\_scores:

Quality:	72.00	Length:	41
Ratio:	3.130 <td>Gaps:</td> <td>2</td>	Gaps:	2
Percent Similarity:	56.098 <td>Percent Identity:</td> <td>39.024</td>	Percent Identity:	39.024

alignment\_block:

US-09-306-689-10 x T42382

Align seg 1/1 to: T42382 from: 1 to: 1068

```

28 GCGAGCGCTGCTCAAGATTGAGC...TAGCGCTGCGTCCGGT..... 69
|||||..... ||| |||||..... |||||
356 G|Y|S|G|Y|S|P|A|P|G|L|E|U|T|P|P|O|V|A|I|Y|L|E|U|G|I|P|P|O|G|I|Y|L|E|A|R 372
70 .....G|C|T|T|A|G|C|C|A|G|A|T|T|G|G|A|G|C|T|A|G|G|C|T|G|G|C 103
372 G|G|Y|L|E|U|A|L|A|T|Y|R|A|G|L|Y|H|S|S|E|R|V|A|L|H|I|S|T|R|P|R|O|H|I|S|S|E|R|S|E|R|P 389
104 CTGGCAGCGGTAGCCAAAGATTGG 126
|||||..... |||
389 roG|Y|T|H|A|S|P|S|E|R|G|Y|C|Y|S|T|R|P 396

```

seq\_name: p1r2:A36443

seq\_documentation\_block:

seminal vesicle secretory protein II precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 05-Nov-1999  
 C:Accession: A36443  
 R:Harris, S.E.; Harris, M.A.; Johnson, C.M.; Bean, M.F.; Dodd, J.G.; Matusik, R.J.; Carr, J. Biol. Chem. 265, 9896-9903, 1990  
 A:Title: Structural characterization of the rat seminal vesicle secretion II protein and  
 A:Reference number: A36443; MUID:90277684  
 A:Accession: A36443  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-414 <HAR>  
 A:Cross-references: GB:J05443; NID:9207114; PIDN:AAA42192.1; PID:9207115

alignment\_scores:  
 Quality: 71.00 Length: 42  
 Ratio: 2.536 Gaps: 0  
 Percent Similarity: 66.667 Percent Identity: 42.857

alignment\_block:  
 US-09-306-689-10 x A36443 ..

Align seg 1/1 to: A36443 from: 1 to: 414

```

10 AGCTACGCGCTGCGCCCTGCGAGCGGTTCTCAAGATTGAGCTACGCGCT 59
|||||..... ||| |||||..... |||||
184 Serheg|Y|G|N|Met|Y|Ser|Ser|G|Y|Ser|In|Val|Y|Ser|Phe|G|Y|G|I 200
60 GCGTCGGGCTGCTTAGCCAGCATTTGAGCTACGCGCTGCGCCCTGGCA 109
|||||..... ||| |||||..... |||
200 nMet|Y|A|L|Ser|G|L|Ser|G|I|n|Met|Y|Ser|Phe|G|Y|G|I|n|A|Y|S|S|E|R|G 217
110 GCGGTAGCCAAAGATTGGAGCTACGCGC 135
|||||..... ||| |||||..... |||
217 InG|Y|G|Y|G|I|n|L|E|U|G|I|n|S|E|R|Y|G|Y|I 225

```

seq\_name: p1r2:T16933

seq\_documentation\_block:

hypothetical protein T24D8.5 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16933  
 R:Martin, J.  
 submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid T24D8.  
 A:Reference number: Z18609  
 A:Accession: T16933  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-171 <HAR>  
 A:Cross-references: EMBL:U04022; NID:g1065529; PID:g1065532; PIDN:AAA81446.1; CESP:T24D8  
 C:Genetics:  
 A:Gene: CESP:T24D8.5

A:Introns: 30/3; 91/2

alignment\_scores:  
 Quality: 70.50 Length: 47  
 Ratio: 2.074 Gaps: 3  
 Percent Similarity: 72.340 Percent Identity: 42.553

alignment\_block:  
 US-09-306-689-10 x T16933 ..

Align seg 1/1 to: T16933 from: 1 to: 171

```

16 GCGCTGCGCCCTGCGAGCGGTTCTCAAGATTGAGC...TAGCGCTGCG 62
|||||..... |||..... |||||
105 G|Y|L|E|U|A|R|P|R|O|G|I|Y|S|A|R|G|S|E|R|M|E|T|A|L|Y|R|G|I|Y|A|R|G|I|n|G|I|P|H|E|R|P 121
63 TCCGGTGCGCTTAGCCAGCATTTGAGC...TAGCGCTGCGCCCTGGCA 109
|||||..... |||..... |||||
121 gProG|Y|S|A|R|G|S|E|R|M|E|T|A|L|Y|R|G|I|Y|A|R|G|I|n|G|I|P|H|E|R|P|roG|Y|L 138
110 GCGGTAGCCAAAGATTGAGC...TAGCGCTGCGTCCGGGT 147
|||..... |||..... |||||
138 Y|S|A|R|G|S|E|R|M|E|T|A|L|Y|R|G|I|Y|A|R|G|I|n|G|I|P|H|E|R|P|roG|Y|L 151

```

seq\_name: p1r4:S33798

seq\_documentation\_block:

FUS/CHOP mutant fusion protein - human  
 N:Alternate names: TLS/CHOP mutant fusion protein  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Nov-1993 #sequence\_revision 30-Nov-1995 #text\_change 20-Apr-2000  
 C:Accession: S33798; S36158  
 R:Crozat, A.; Aman, P.; Mandahl, N.; Ron, D.  
 Nature 363, 640-644, 1993  
 A:Title: Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma.  
 A:Reference number: S33798; MUID:93288139  
 A:Accession: S33798  
 A:Molecule type: mRNA  
 A:Residues: 1-462 <CRO>  
 A:Cross-references: GB:S62138; NID:9386158; PIDN:AA827103.1; PID:9386159  
 R:Rabblits, T.H.; Forster, A.; Larson, R.; Nathan, P.  
 Nature Genet. 4, 175-180, 1993  
 A:Title: Fusion of the dominant negative transcription regulator CHOP with a novel ge  
 A:Reference number: S36157  
 A:Accession: S36158  
 A:Molecule type: mRNA  
 A:Residues: 1-462 <RAB>  
 A:Cross-references: EMBL:X71427; NID:g395919; PIDN:CAA50558.1; PID:e1373214; PID:g4421  
 C:Comment: This sequence is the chimeric product of a translocation mutation.  
 C:Genetics:  
 A:Gene: GADD153/FUS  
 A:Map position: 12q43/16p11  
 A:Note: TLS is a synonym for GDB:FUS  
 C:Keywords: fusion protein

alignment\_scores:  
 Quality: 69.00 Length: 45  
 Ratio: 2.464 Gaps: 1  
 Percent Similarity: 62.222 Percent Identity: 40.000

alignment\_block:  
 US-09-306-689-10 x S33798 ..

Align seg 1/1 to: S33798 from: 1 to: 462

```

13 TAGCGCTGCGCCCTGCGAGCGGTTCTCAAGATTGAGCTACGCGCTGCG 62
|||||..... |||..... |||||
100 T|Y|R|I|Y|G|I|n|P|R|O|A|I|A|P|R|O|S|E|R|T|H|S|E|R|Y|S|E|R|Y|G|Y|S|E|R|S|E 116
63 TCCGGTGCGCTTAGC.....CAGCATTTGAGCTACGCGCTGC 100
|||..... |||..... |||||

```



114 TAGCCAGAT 123  
 |||  
 33 YLYSArgAsp 36

seq\_name: p1r2:T00975

seq\_documentation\_block:  
 hypothetical protein T9J22.9 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 14-May-1999  
 C:Accession: T00975  
 R:Roundslay, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
 submitted to the EMBL Data Library, April 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC T9J22 genomic sequence.  
 A:Reference number: 214153  
 A:Accession: T00975  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-705 <R0U>  
 A:Cross-References: EMBL:AC002505; NID:92739359; PID:92739367  
 A:Experimental source: cultivar Columbia  
 A:Genetics:  
 A:Map position: 2  
 A:Introns: 395/2; 450/3; 487/3; 591/1; 602/2  
 A:Note: T9J22.9

alignment\_scores:  
 Quality: 68.50 Length: 44  
 Ratio: 2.635 Gaps: 1  
 Percent Similarity: 59.091 Percent Identity: 31.818

alignment\_block:  
 US-09-306-689-10 x T00975 ..

Align seg 1/1 to: T00975 from: 1 to: 705

4 CATGTGAGCTACGCGCTGCGCCCTGCGAGCGCTCTCAAGATTGAGCTA 53  
 |||  
 130 HistripleutropglyArglyShisglyTyrGlyIululsarglyTyrAlaAs 146  
 54 CGGCTGCTGCTGCGGCTCTGAGCCAGCATGTGAGCTACGCGCTGCGCC 103  
 |||  
 146 ngly.....AspGlyTyrGlnGlyAsnTrpIysAlaAsnIeuGlnA 160  
 104 CTGGCAGCGGTAGCCAGATTGAGCTACGCGCTACGGC 135  
 |||  
 160 spGlyAsnGlyArgTyrValTrpSerAspGly 170

seq\_name: p1r2:A43900

seq\_documentation\_block:  
 spicule matrix protein LSM34 precursor - sea urchin (Lytechinus pictus)  
 C:Species: Lytechinus pictus (painted urchin)  
 C:Date: 10-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
 C:Accession: A43900  
 R:Livingston, B.T.; Shaw, R.; Bailey, A.; Willt, F.  
 Dev. Biol. 148, 473-480, 1991  
 A:Title: Characterization of a cDNA encoding a protein involved in formation of the skele  
 A:Reference number: A43900; M0ID:92077276  
 A:Accession: A43900  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-335 <LIIV>  
 A:Cross-References: GB:X59616; GB:S68975; NID:9288149; PID:9288150  
 A:Note: sequence extracted from NCBI backbone (NCBIN:68975, NCBIIP:68978)  
 C:Keywords: tandem repeat

alignment\_scores:  
 Quality: 68.00 Length: 46  
 Ratio: 2.000 Gaps: 2  
 Percent Similarity: 73.913 Percent Identity: 39.130

alignment\_block:  
 US-09-306-689-10 x A43900 ..

Align seg 1/1 to: A43900 from: 1 to: 335

16 GGCCTGCGCCCTGCGAGCGCTCTCAAGATTGAGCTACGCGCTGCGT... 63  
 |||  
 204 GlyGlnGlnProGlyPheGlyGlnGlnProGlyPheGlyGlyArgGly 220  
 64 .CGGCTGCTCTAGCCAGCATGTGAGCTACGCGCTGCGCTGCGCAGCG 112  
 |||  
 220 nProGlyPheGlyGlyGlnGlnProGlyPheGlyGlnGlnProGlyPheG 237  
 113 GTAGCCAGATTGAGCTACGCGCTGCGT...CGGGT 147  
 |||  
 237 LyGlyArgGlnProGlyPheGlyGlyArgGlnProGly 249



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:36 ; Search time 47.48 Seconds  
(without alignments)  
6.051 Million cell updates/sec

Title: US-09-306-689-5

Perfect score: 101

Sequence: 1 XHMSYGLRPPPPPPC 16

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/PCYUS.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	100	99.0	16	1	US-08-188-223-2	Sequence 2, Appl
2	100	99.0	16	1	US-08-188-223-2	Sequence 2, Appl
3	83.5	82.7	17	1	US-08-188-223-6	Sequence 6, Appl
4	83.5	82.7	17	1	US-08-188-223-6	Sequence 6, Appl
5	63	62.4	15	1	US-08-453-588-25	Sequence 25, Appl
6	63	62.4	15	1	US-08-453-588-25	Sequence 25, Appl
7	63	62.4	15	1	US-08-453-588-25	Sequence 25, Appl
8	63	62.4	15	1	US-08-453-588-25	Sequence 25, Appl
9	62	61.4	16	3	US-08-521-079-26	Sequence 26, Appl
10	62	61.4	16	3	US-08-521-079-26	Sequence 26, Appl
11	62	61.4	16	3	US-08-458-814-7	Sequence 7, Appl
12	58	57.4	10	1	US-07-714-540-9	Sequence 9, Appl
13	58	57.4	10	1	US-07-690-983D-2	Sequence 2, Appl
14	58	57.4	10	1	US-07-690-983D-2	Sequence 32, Appl
15	58	57.4	10	1	US-08-103-022-1	Sequence 1, Appl
16	58	57.4	10	1	US-08-184-935-6	Sequence 6, Appl
17	58	57.4	10	1	US-08-343-883-1	Sequence 1, Appl
18	58	57.4	10	1	US-08-000-931-5	Sequence 5, Appl
19	58	57.4	10	1	US-08-428-488-22	Sequence 22, Appl
20	58	57.4	10	1	US-08-341-219-11	Sequence 11, Appl
21	58	57.4	10	1	US-08-453-588-4	Sequence 4, Appl
22	58	57.4	10	1	US-08-453-588-4	Sequence 4, Appl
23	58	57.4	10	1	US-08-453-588-6	Sequence 6, Appl
24	58	57.4	10	1	US-08-453-588-8	Sequence 8, Appl
25	58	57.4	10	1	US-08-453-588-10	Sequence 10, Appl
26	58	57.4	10	1	US-08-453-588-12	Sequence 12, Appl
27	58	57.4	10	1	US-08-453-588-14	Sequence 14, Appl
28	58	57.4	10	1	US-08-453-588-16	Sequence 16, Appl

29	58	57.4	10	1	US-08-453-588-22	Sequence 22, Appl
30	58	57.4	10	1	US-08-188-223-3	Sequence 3, Appl
31	58	57.4	10	1	US-08-406-935-5	Sequence 1, Appl
32	58	57.4	10	1	US-08-591-917-1	Sequence 1, Appl
33	58	57.4	10	1	US-08-387-156-2	Sequence 2, Appl
34	58	57.4	10	1	US-08-474-555-1	Sequence 1, Appl
35	58	57.4	10	1	US-08-446-692-1	Sequence 1, Appl
36	58	57.4	10	1	US-08-242-678D-1	Sequence 1, Appl
37	58	57.4	10	2	US-08-796-598-6	Sequence 6, Appl
38	58	57.4	10	2	US-08-694-865-2	Sequence 2, Appl
39	58	57.4	10	2	US-08-694-865-18	Sequence 18, Appl
40	58	57.4	10	2	US-08-488-351A-1	Sequence 1, Appl
41	58	57.4	10	2	US-08-480-494B-1	Sequence 1, Appl
42	58	57.4	10	2	US-08-447-175A-6	Sequence 6, Appl
43	58	57.4	10	2	US-08-878-748-2	Sequence 2, Appl
44	58	57.4	10	3	US-08-521-079-2	Sequence 2, Appl
45	58	57.4	10	5	5492893-1	Patent No. 5492893

## ALIGNMENTS

RESULT 1  
US-08-188-223-2  
Sequence 2, Application US/08188223  
Patent No. 5686506  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Scibienski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESS: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,223  
FILING DATE: 27-JAN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: YES  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note= "immunomic"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11..16  
OTHER INFORMATION: /note= "spacer"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1

OTHER INFORMATION: /label= pglu  
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"  
US-08-188-223-2

Query Match 99.0%; Score 100; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 5.5e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 HWSYGLRGRPPPC 16  
Db 2 HWSYGLRGRPPPC 16

RESULT 2  
US-08-968-466-2  
Sequence 2, Application US/08968466  
Patent No. 613720  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Sciabenski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968,466  
FILING DATE: 27-JAN-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11..16  
OTHER INFORMATION: /note= "spacer"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label= pglu  
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"  
US-08-968-466-2

Query Match 99.0%; Score 100; DB 3; Length 16;  
Best Local Similarity 100.0%; Pred. No. 5.5e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRGRPPPC 16  
Db 2 HWSYGLRGRPPPC 16

RESULT 3  
US-08-188-223-6  
Sequence 6, Application US/08188223  
Patent No. 5688506  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Sciabenski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,223  
FILING DATE: 27-JAN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label= pglu  
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11..17  
OTHER INFORMATION: /note= "spacer"  
US-08-188-223-6

Query Match 82.7%; Score 83.5; DB 1; Length 17;  
Best Local Similarity 87.5%; Pred. No. 6.8e-05;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 HWSYGLRPG-RPPPC 16  
Db 2 HWSYGLRPGSSPPPC 17  
RESULT 4  
US-08-968-466-6

Sequence 6, Application US/08968466  
Patent No. 6132720  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968,466  
FILING DATE: 27-JAN-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-354-8113  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label= pglu  
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11..17  
OTHER INFORMATION: /note= "spacer"  
US-08-968-466-6

Query Match 82.7%; Score 83.5; DB 3; Length 17;  
Best Local Similarity 87.5%; Pred. No. 6.8e-05;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 HWSYGLRPG-RPPPPC 16  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 5  
US-08-453-588-25  
Sequence 25, Application US/08453588  
Patent No. 5684145  
GENERAL INFORMATION:  
APPLICANT: Anna van der Zee, Irma Marianne van Die,  
APPLICANT: Willem Pieter Martin Hoekstra,  
APPLICANT: Josephus Theodorus Gielen,  
TITLE OF INVENTION: Carrier system against GnRH  
NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo No. 5684145el Patent Department  
STREET: 1300 Piccard Drive, Suite 206  
CITY: Rockville  
STATE: Maryland  
COUNTRY: U.S.A.  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,588  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/078,661  
FILING DATE: 16-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Maity E. Gormley  
REGISTRATION NUMBER: 34,409  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 258-5200  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
US-08-453-588-25

Query Match 62.4%; Score 63; DB 1; Length 15;  
Best Local Similarity 90.9%; Pred. No. 0.023;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGRP 12  
DB 4 HWSYGLRPGDP 14

RESULT 6  
US-08-521-079-25  
Sequence 25, Application US/08521079  
Patent No. 6019983  
GENERAL INFORMATION:  
APPLICANT: Anna van der Zee, Irma Marianne van Die,  
APPLICANT: Willem Pieter Martin Hoekstra,  
APPLICANT: Josephus Theodorus Gielen,  
TITLE OF INVENTION: Carrier system against GnRH  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo Pharma  
STREET: 1330 Piccard Drive  
CITY: Rockville  
STATE: Maryland  
COUNTRY: U.S.A.  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/521,079  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/078,661  
FILING DATE: 16-JUN-1993  
APPLICATION NUMBER: EPA No. 6019983 92.201.775.1  
FILING DATE: 18-JUN-1992

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: William M. Blackstone  
REGISTRATION NUMBER: 29,772  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 258-5200  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-521-079-25

Query Match 62.4%; Score 63; DB 3; Length 15;  
Best Local Similarity 90.9%; Pred. No. 0.023;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRGRP 12  
Db 4 HWSYGLRGRP 14

RESULT 7  
US-08-453-588-26  
Sequence 26, Application US/08453588  
Patent No. 5684145  
GENERAL INFORMATION:  
APPLICANT: Anna van der Zee, Irma Marianne van Die,  
APPLICANT: Willem Pieter Martin Hoekstra,  
APPLICANT: Josephus Theodorus Gielen.  
TITLE OF INVENTION: Carrier system against GNRH  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo No. 5684145el Patent Department  
STREET: 1300 Piccard Drive, Suite 206  
CITY: Rockville  
STATE: Maryland  
COUNTRY: U.S.A.  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,588  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/078,661  
FILING DATE: 16-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary E. Gormley  
REGISTRATION NUMBER: 34,409  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 258-5200  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-588-26

Query Match 62.4%; Score 63; DB 1; Length 16;  
Best Local Similarity 90.9%; Pred. No. 0.024;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRGRP 12

Db 5 HWSYGLRGRP 15

RESULT 8  
US-08-521-079-26  
Sequence 26, Application US/08521079  
Patent No. 6019983  
GENERAL INFORMATION:  
APPLICANT: Anna van der Zee, Irma Marianne van Die,  
APPLICANT: Willem Pieter Martin Hoekstra,  
APPLICANT: Josephus Theodorus Gielen.  
TITLE OF INVENTION: Carrier system against GNRH  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo Pharma  
STREET: 1330 Piccard Drive  
CITY: Rockville  
STATE: Maryland  
COUNTRY: U.S.A.  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/521,079  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/078,661  
FILING DATE: 16-JUN-1993  
APPLICATION NUMBER: EPA No. 6019983 92.201.775.1  
FILING DATE: 18-JUN-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: William M. Blackstone  
REGISTRATION NUMBER: 29,772  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 258-5200  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-521-079-26

Query Match 62.4%; Score 63; DB 3; Length 16;  
Best Local Similarity 90.9%; Pred. No. 0.024;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRGRP 12  
Db 5 HWSYGLRGRP 15

RESULT 9  
US-08-458-814-6  
Sequence 6, Application US/08458814  
Patent No. 6103243  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J  
APPLICANT: DE AIZPURA, Henry J  
APPLICANT: HOME, Peter  
APPLICANT: RAND, Keith N  
TITLE OF INVENTION: ORAL VACCINES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,814  
FILING DATE: 02-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/327,822  
FILING DATE: 18-OCT-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER: PCT/AU86/00135  
FILING DATE: 14-MAY-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PH3104  
FILING DATE: 25-OCT-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PH0566  
FILING DATE: 15-MAY-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 60042/155/BIANU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 52 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-458-814-6

Query Match 61.4%; Score 62; DB 3; Length 52;  
Best Local Similarity 83.3%; Pred. No. 0.096;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRP 13  
DB 37 HWSYGLRPGDP 48

RESULT 10  
US-08-458-814-7  
Sequence 7, Application US/08458814  
Patent No. 6103243  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J  
APPLICANT: DE AIZPURA, Henry J  
APPLICANT: HOME, Peter  
APPLICANT: RAND, Keith N  
TITLE OF INVENTION: ORAL VACCINES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,814  
FILING DATE: 02-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/327,822  
FILING DATE: 18-OCT-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER: PCT/AU86/00135  
FILING DATE: 14-MAY-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PH3104  
FILING DATE: 25-OCT-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PH0566  
FILING DATE: 15-MAY-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 60042/155/BIANU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-458-814-7

Query Match 61.4%; Score 62; DB 3; Length 55;  
Best Local Similarity 83.3%; Pred. No. 0.1;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRP 13  
DB 39 HWSYGLRPGDP 50

RESULT 11  
US-07-714-540-9  
Sequence 9, Application US/07714540  
Patent No. 5262521  
GENERAL INFORMATION:  
APPLICANT: Almquist, Ronald G.  
APPLICANT: Toll, Lawrence  
TITLE OF INVENTION: ISOLATED ATRIAL PEPTIDE-DEGRADING  
TITLE OF INVENTION: ENZYME AND NOVEL COMPOUNDS USEFUL AS INHIBITORS THEREOF  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Irell & Manella  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: California  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/714,540  
FILING DATE: 19910607  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Reed, Dianne E.

REGISTRATION NUMBER: 31,292  
REFERENCE/DOCKET NUMBER: 8500-0135.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-327-7250  
TELEFAX: 415-327-2951  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-714-540-9

Query Match 57.4%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.067;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 12  
US-07-690-983D-2  
Sequence 2, Application US/07690983D  
Patent No. 5403586  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J.  
APPLICANT: STEWART, Andrew G.  
APPLICANT: TSONIS, Con G.  
TITLE OF INVENTION: FUSION PROTEINS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/690,983D  
FILING DATE: 25-JUN-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU90/00373  
FILING DATE: 24-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/148 CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-07-690-983D-2

Query Match 57.4%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.067;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 13  
US-07-690-983D-32  
Sequence 32, Application US/07690983D  
Patent No. 5403586  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J.  
APPLICANT: STEWART, Andrew G.  
APPLICANT: TSONIS, Con G.  
TITLE OF INVENTION: FUSION PROTEINS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/690,983D  
FILING DATE: 25-JUN-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU90/00373  
FILING DATE: 24-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/148 CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-690-983D-32

Query Match 57.4%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.067;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 14  
US-08-103-022-1  
Sequence 1, Application US/08103022  
Patent No. 5413990  
GENERAL INFORMATION:  
APPLICANT: Haviv, Fortuna  
APPLICANT: Fitzpatrick, Timothy D.  
APPLICANT: Swenson, Rolf E.  
APPLICANT: Nichols, Charles J.  
APPLICANT: Mott, Nicholas A.  
TITLE OF INVENTION: N-Terminus Modified Analogs of LHRH  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377  
STREET: Abbott Laboratories, One Abbott Park Road

CITY: NO. 5413990Lh Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/103.022  
FILING DATE: 05-OCT-1993  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Janssen, Jerry F.  
REGISTRATION NUMBER: 29,175  
REFERENCE/DOCKET NUMBER: 5389.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708) 938-7742  
TELEFAX: (708) 938-2623  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /note= "Xaa at position 1 is a  
5-oxo-prolyl aminoacyl residue."  
US-08-103-022-1

Query Match 57.4%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.067;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
|||||  
DB 2 HWSYGLRPG 10

RESULT 15  
US-08-184-935-6  
Sequence 6, Application US/08184935  
Patent No. 5476770  
GENERAL INFORMATION:  
APPLICANT: PRADILLES, PHILIPPE  
TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN  
TITLE OF INVENTION: OR HAPTEN  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/184.935  
FILING DATE: 24-JAN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5476770man F.  
REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 846-286-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /note= "C-terminal amide"  
US-08-184-935-6

Query Match 57.4%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.067;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
|||||  
DB 2 HWSYGLRPG 10

Search completed: March 2, 2001, 10:54:36  
Job time: 601 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:00:48 ; Search time 97.13 Seconds  
(without alignments)  
19.307 Million cell updates/sec

Title: US-09-306-689-5  
Perfect score: 101  
Sequence: 1 XHMYGLRPGRPPEPC 16

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues  
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_15:\*  
2: sp.archaea:\*  
3: sp.bacteria:\*  
4: sp.fungi:\*  
5: sp.human:\*  
6: sp.invertebrate:\*  
7: sp.mammal:\*  
8: sp.mhc:\*  
9: sp.organelle:\*  
10: sp.plant:\*  
11: sp.protist:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	57.4	91	13	Q9PRH0 anguilla ja
2	55	54.5	969	5	Q17795 caenorhabdi
3	54	53.5	381	5	Q9YWD2 drosophila
4	53	52.5	112	13	Q88020 chimpanzee
5	52	51.5	87	12	Q9Y126 sparus aura
6	52	51.5	95	13	Q73812 morone saxa
7	52	51.5	99	13	Q91A10 dicentrarch
8	52	51.5	112	12	Q04273 chimpanzee
9	52	51.5	112	12	Q88050 chimpanzee
10	52	51.5	112	12	Q88057 chimpanzee
11	52	51.5	112	12	Q88058 chimpanzee
12	52	51.5	112	12	Q88064 chimpanzee
13	52	51.5	112	12	Q88068 chimpanzee
14	52	51.5	112	12	Q88125 chimpanzee
15	52	51.5	112	12	Q88131 chimpanzee
16	52	51.5	112	12	Q88137 chimpanzee
17	52	51.5	112	12	Q88145 chimpanzee
18	52	51.5	112	12	Q89827 chimpanzee
19	52	51.5	112	12	Q07393 chimpanzee

20	52	51.5	112	12	Q85604 human t-cell
21	52	51.5	112	12	Q89552 chimpanzee
22	51	50.5	263	13	Q9PT52 agkistrodon
23	51	50.5	2061	5	Q9YU93 drosophila
24	50	49.5	90	13	Q91A02 human immun
25	50	49.5	112	12	Q76623 human immun
26	50	49.5	112	12	P88144 human immun
27	50	49.5	112	12	Q90159 human immun
28	50	49.5	112	12	Q90158 human immun
29	50	49.5	112	12	Q90157 human immun
30	50	49.5	112	12	Q90156 human immun
31	50	49.5	112	12	Q9PX03 human immun
32	50	49.5	211	5	Q45098 human immun
33	50	49.5	480	6	Q06018 ovis aries
34	49.5	48.0	423	12	Q12272 simian t-ce
35	49	48.5	373	4	Q9NP08 homo sapien
36	49	48.5	645	4	Q9P282 homo sapien
37	49	48.5	754	5	Q9NSW3 caenorhabdi
38	49	48.5	2129	5	Q9Q1R8 homo sapien
39	48.5	48.0	1049	4	Q94957 homo sapien
40	48	47.5	329	12	Q09798 human herpe
41	48	47.5	598	3	Q9P6X0 neurospora
42	48	47.5	629	5	Q07997 chironomus
43	48	47.5	635	12	Q69087 human herpe
44	48	47.5	1051	5	Q9Y150 drosophila
45	48	47.5	1579	5	Q9Y451 drosophila

## ALIGNMENTS

RESULT 1  
ID Q9PRH0 PRELIMINARY: PRT: 91 AA.

AC Q9PRH0;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DI 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE PREPRO-MGNRH PRECURSOR.  
OS Anguilla japonica (Japanese eel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;  
OC Anguillidae; Anguilla.  
OX NCBI\_Taxid=7937;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA Kubo K., Suetake H., Aida K.;  
RT "Expression of two gonadotropin-releasing hormone (GNRH) precursor genes in various tissues of the Japanese eel and evolution of GNRH.";  
RL Zool. Sci. 16:471-478(1999).  
RN (2)  
RP SEQUENCE FROM N.A.  
RA Kubo K., Suetake H., Aida K.;  
RT "A splicing variant for the prepro-mammalian gonadotropin-releasing hormone (prepro-mgnrh) mRNA is present in the brain and various peripheral tissues of the Japanese eel.";  
RL Zool. Sci. 16:645-651(1999).  
DR EMBL; AB026989; BAA82608.1; -  
DR EMBL; AB026991; BAA83597.1; -  
DR INTERPRO; IPR02012; -  
DR PFAM; PF00446; GNRH.1.  
DR PROSITE; PS00473; GNRH.1.  
KW Signal.  
FT SIGNAL. 1 22  
FT CHAIN 23 32  
FT CHAIN 33 91  
SQ SEQUENCE 91 AA: 9893 MW: 84159DC08434A7B CRC64;

Query Match 57.4%; Score 58; DB 13; Length 91;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||  
 Db 24 HWSYGLRPG 32

## RESULT 2

Q17795 PRELIMINARY; PRT; 969 AA.

ID Q17795  
 AC Q17795;  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)  
 DE SIMILARITY TO WISCOTT-ALDRICH SYNDROME PROTEINS.  
 GN CG14234.  
 OS Eukaryota: Metazoa: Nematoda: Chromadorea: Rhabditidae; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OC NCBI\_Taxid=6239;  
 RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Faveille A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Keshaw J., Kirsten T., Laister N., Latreille P.,  
 RA Lighting J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).

RL [2]

SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Hawkins J.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U58751; AAB00657.1; -.

DR INTERPRO: IPR000095; -.

DR INTERPRO: IPR000697; -.

DR INTERPRO: IPR001124; -.

DR PFAM: PF00568; WH1; 1.

DR PFAM: PF00786; PBD; 1.

DR PFAM: PF02205; WH2; 2.

SEQUENCE 969 AA; 105790 MW; BEZD11SD0C620HE8 CRC64;

Query Match 54.5%; Score 55; DB 5; Length 969;

Best Local Similarity 75.0%; Pred. No. 7;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 SYGLRPGRPP 15  
 |||||  
 Db 727 SHGLAPARPPP 738

## RESULT 3

Q9VMD2 PRELIMINARY; PRT; 381 AA.

AC Q9VMD2;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)

DE Chimpanzee immunodeficiency virus (SIVcpz) (CIV).

OS Chimpanzee immunodeficiency virus (SIVcpz) (CIV).

CG14234 PROTEIN.

GN CG14234.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta;

OC Pterygota: Neoptera: Endopterygota: Diptera; Brachycera: Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OC NCBI\_Taxid=7227;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,

RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doug L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman J.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,

RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,

RT "The genome sequence of Drosophila melanogaster";

RL Science 287:2185-2195(2000).

DR EMBL: AE003513; AAF49014.1; -.

DR FLYBASE: FBgn0031065; CG14234.

SEQUENCE 381 AA; 41732 MW; 77317DAE40A863 CRC64;

Query Match 53.5%; Score 54; DB 5; Length 381;

Best Local Similarity 53.3%; Pred. No. 4;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRPP 16  
 |||||  
 Db 205 NWIGMRHPSPPC 219

## RESULT 4

Q88020 PRELIMINARY; PRT; 112 AA.

AC Q88020;  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)

DE (M316ZA) DNA, COMPLETE CDS'S.

OS Chimpanzee immunodeficiency virus (SIVcpz) (CIV).

OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID-11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94016835; PubMed-8411355;  
 RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
 RT "Analysis of simian immunodeficiency virus sequence variation in  
 tissues of rhesus macaques with simian AIDS."  
 RL J. Virol. 67:6522-6534(1993).  
 DR EMBL: L22807; AAA47688.1; -.  
 DR INTERPRO: IPR000012; -.  
 DR PFAM: PF00522; VPR; 1.  
 DR PRINTS: PR00444; HIVPRVFX.  
 SQ SEQUENCE 112 AA; 12945 MW; A933BF982AFD7BE8 CRC64;

Query Match 52.5%; Score 53; DB 12; Length 112;  
 Best Local Similarity 64.3%; Pred. No. 1.7;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 OY 2 HWSYGLRPGPPP 15  
 Db 94 HGAEGWRPGPPPP 107

RESULT 5  
 OY126 PRELIMINARY; PRT; 87 AA.  
 AC 09Y126;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)  
 DE (FRAGMENT)  
 OS Sparus aurata (Gilthead sea bream).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC NCBI\_TaxID-8175;  
 OX NCBI\_TaxID-8175;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Ovary;  
 RA Nabissi M.;  
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.  
 CC - FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC - FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
 CC EMBL: AF046801; AAD02427.1; -.  
 DR INTERPRO: IPR002012; -.  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KW Hormone; Amidation.  
 FT NON\_TER 1 87  
 FT NON\_TER 87  
 SQ SEQUENCE 87 AA; 9871 MW; 0D2463533D96782A CRC64;

Query Match 51.5%; Score 52; DB 13; Length 87;  
 Best Local Similarity 88.9%; Pred. No. 1.9;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 HWSYGLRPG 10  
 Db 22 HWSYGLSPG 30

RESULT 6  
 OY3812 PRELIMINARY; PRT; 95 AA.  
 AC 073812;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).  
 OS Morone saxatilis (Striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Moronidae; Morone.  
 OX NCBI\_TaxID-34816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chow M.M., Kight K.E., Gotthilf Y., Alok D., Zohar Y.;  
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.  
 CC - FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC - FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
 CC EMBL: AF056314; AAD03817.1; -.  
 DR INTERPRO: IPR002012; -.  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 DR PRODOM: PD005581; -; 1.  
 KW Hormone; Amidation.  
 SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match 51.5%; Score 52; DB 13; Length 95;  
 Best Local Similarity 88.9%; Pred. No. 2.1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 HWSYGLRPG 10  
 Db 24 HWSYGLSPG 32

RESULT 7  
 OY1A10 PRELIMINARY; PRT; 99 AA.  
 AC 091A10;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)  
 DE GONADOTROPIN-RELEASING HORMONE SEABREAM ISOFORM.  
 OS Dicentrarchus labrax (European sea bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Moronidae; Dicentrarchus.  
 OX NCBI\_TaxID-13489;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanuy S.,  
 RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;  
 RT "Differential expression of three different prepro-GNRH  
 RT (Gonadotrophin-releasing hormone) messengers in the brain of the  
 RT European sea bass (Dicentrarchus labrax)."  
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Zmora N., Zohar Y., Elizur A.;  
 RT "3 GNRH form in the seabass Dicentrarchus labrax";  
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF224279; AAF62898.1; -.  
 SQ SEQUENCE 99 AA; 10758 MW; ECA8AEC93CC02904 CRC64;

Query Match 51.5%; Score 52; DB 13; Length 99;  
 Best Local Similarity 88.9%; Pred. No. 2.1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 HWSYGLRPG 10  
 Db 28 HWSYGLSPG 36

RESULT 8  
ID 004273 PRELIMINARY; PRT; 112 AA.  
AC 004273;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VPX PROTEIN (X ORF PROTEIN).  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Benveniste R.E., Heidecker G., Greenwood J., Gonda M.A.;  
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.  
CC -1- MISCELLANEOUS: THIS IS A CAPTIVE MACAQUE NEMESTRINA ISOLATE.  
DR EMBL: M32741; AAA91933.1; -  
DR INTERPRO: IPR00012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVPX.  
KW AIDS.  
SQ SEQUENCE 112 AA; 12904 MW; 7E1564FA5AFD7BEA CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;  
Best Local Similarity 64.3%; Pred. No. 2.4;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRGRPPPP 15  
ID 94 HGAGGWRGPPPP 107  
Db

RESULT 9  
ID 088050 PRELIMINARY; PRT; 112 AA.  
AC 088050;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE (MM316X) DNA, COMPLETE CDS'S.  
GN VPR.  
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94016835; PubMed=8411355;  
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
RT "Analysis of simian immunodeficiency virus sequence variation in  
RT tissues of rhesus macaques with simian AIDS."  
RL J. Virol. 67:6522-6534(1993).  
DR EMBL: L22804; AAA47679.1; -  
DR INTERPRO: IPR00012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVPX.  
SQ SEQUENCE 112 AA; 12872 MW; 44BD094EF1955B60 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;  
Best Local Similarity 64.3%; Pred. No. 2.4;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRGRPPPP 15  
ID 94 HGAGGWRGPPPP 107  
Db

RESULT 10  
ID 088057

ID 088057 PRELIMINARY; PRT; 112 AA.  
AC 088057;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE (MM316K) DNA, COMPLETE CDS'S.  
GN VPR.  
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94016835; PubMed=8411355;  
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
RT "Analysis of simian immunodeficiency virus sequence variation in  
RT tissues of rhesus macaques with simian AIDS."  
RL J. Virol. 67:6522-6534(1993).  
DR EMBL: L22791; AAA72399.1; -  
DR INTERPRO: IPR00012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVPX.  
SQ SEQUENCE 112 AA; 12842 MW; BDB78F982AFD7BF2 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;  
Best Local Similarity 64.3%; Pred. No. 2.4;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRGRPPPP 15  
ID 94 HGAGGWRGPPPP 107  
Db

RESULT 11  
ID 088058 PRELIMINARY; PRT; 112 AA.  
AC 088058;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE (MM316L) DNA, COMPLETE CDS'S.  
GN VPR.  
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94016835; PubMed=8411355;  
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
RT "Analysis of simian immunodeficiency virus sequence variation in  
RT tissues of rhesus macaques with simian AIDS."  
RL J. Virol. 67:6522-6534(1993).  
DR EMBL: L22792; AAA47643.1; -  
DR INTERPRO: IPR00012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVPX.  
SQ SEQUENCE 112 AA; 12846 MW; 7FC007A3BAFD7BEA CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;  
Best Local Similarity 64.3%; Pred. No. 2.4;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRGRPPPP 15  
ID 94 HGAGGWRGPPPP 107  
Db

RESULT 12  
ID 088064 PRELIMINARY; PRT; 112 AA.  
AC 088064;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE (MM3160) DNA, COMPLETE CDS'S.  
 GN VPX.  
 OS Chimpanzee immunodeficiency virus (STV(cpz)) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94016835; PubMed=8411355;  
 RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
 RT "Analysis of simian immunodeficiency virus sequence variation in  
 RT tissues of rhesus macaques with simian AIDS.";  
 RL J. Virol. 67:6522-6534(1993).  
 DR EMBL: L22795: AAA47652.1; -;  
 DR INTERPRO: IPR000012; -;  
 DR PFAM: PF00522; VPR: 1.  
 DR PRINTS: PR00444; HIVPRVPX.  
 SQ SEQUENCE 112 AA; 12935 MW; D32070A4DAC1CF5 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;  
 Best Local Similarity 64.3%; Pred. No. 2.4;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HMSGRLRPGRP 15  
 DB 94 HGAGGWRPGRPP 107

RESULT 13  
 ID 088068 PRELIMINARY; PRT: 112 AA.  
 AC 088068;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE (MM3160) DNA, COMPLETE CDS'S.  
 GN VPX.  
 OS Chimpanzee immunodeficiency virus (STV(cpz)) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94016835; PubMed=8411355;  
 RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
 RT "Analysis of simian immunodeficiency virus sequence variation in  
 RT tissues of rhesus macaques with simian AIDS.";  
 RL J. Virol. 67:6522-6534(1993).  
 DR EMBL: L22797: AAA47658.1; -;  
 DR INTERPRO: IPR000012; -;  
 DR PFAM: PF00522; VPR: 1.  
 DR PRINTS: PR00444; HIVPRVPX.  
 SQ SEQUENCE 112 AA; 12841 MW; 8B82EF96CA137BF0 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;  
 Best Local Similarity 64.3%; Pred. No. 2.4;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HMSGRLRPGRP 15  
 DB 94 HGAGGWRPGRPP 107

RESULT 14  
 ID 088125 PRELIMINARY; PRT: 112 AA.  
 AC 088125;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE VPX PROTEIN.

GN VPX.  
 OS Chimpanzee immunodeficiency virus (STV(cpz)) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90272009; PubMed=1971917;  
 RA Demhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;  
 RT "Sequence analysis and acute pathogenicity of molecularly cloned  
 RT SIVSM-PBj14.";  
 RL Nature 345:636-640(1990).  
 DR EMBL: L03296; AAA47763.1; -;  
 DR INTERPRO: IPR000012; -;  
 DR PFAM: PF00522; VPR: 1.  
 DR PRINTS: PR00444; HIVPRVPX.  
 SQ SEQUENCE 112 AA; 12877 MW; BEADC7FF6CED2EAC CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;  
 Best Local Similarity 64.3%; Pred. No. 2.4;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HMSGRLRPGRP 15  
 DB 94 HGAGGWRPGRPP 107

RESULT 15  
 ID 088131 PRELIMINARY; PRT: 112 AA.  
 AC 088131;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE VPX PROTEIN.  
 GN VPX.  
 OS Chimpanzee immunodeficiency virus (STV(cpz)) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90272009; PubMed=1971917;  
 RA Demhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;  
 RT "Sequence analysis and acute pathogenicity of molecularly cloned  
 RT SIVSM-PBj14.";  
 RL Nature 345:636-640(1990).  
 DR EMBL: L03297; AAA47771.1; -;  
 DR INTERPRO: IPR000012; -;  
 DR PFAM: PF00522; VPR: 1.  
 DR PRINTS: PR00444; HIVPRVPX.  
 SQ SEQUENCE 112 AA; 12851 MW; 9C3EC7FF6CED3D14 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;  
 Best Local Similarity 64.3%; Pred. No. 2.4;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HMSGRLRPGRP 15  
 DB 94 HGAGGWRPGRPP 107

Search completed: March 2, 2001, 11:00:49  
 Job time: 459 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:01:20 ; Search time 32.03 Seconds

(without alignments)  
16.132 Million cell updates/sec

Title: US-09-306-689-5  
Perfect score: 101  
Sequence: 1 XHMSYGLRPGRRPPPC 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues  
Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	57.4	61	1 GONI_SHEEP	Q28588 ovis aries
2	58	57.4	63	1 GONI_MESAP	O09163 mesocricetu
3	58	57.4	67	1 GONI_MACMU	P55247 macaca mula
4	58	57.4	89	1 GONI_XENLA	P45656 xenopus lae
5	58	57.4	90	1 GONI_MOUSE	P13662 mus musculu
6	58	57.4	91	1 GONI_PIG	P49921 sus scrofa
7	58	57.4	92	1 GONI_HUMAN	P01148 homo sapien
8	58	57.4	92	1 GONI_RAT	P07490 rattus norv
9	58	57.4	92	1 GONI_TUPGB	O95335 tupiaa gils
10	54	53.5	10	1 GONI_ALAMI	P37041 alligator m
11	54	53.5	92	1 GONI_CHICK	P37042 gallus gall
12	52	51.5	94	1 GONI_HAPBU	P51918 haplochromi
13	52	51.5	95	1 GONI_PAGMA	P70074 pagrus majo
14	52	51.5	95	1 GONI_SPAUV	P51919 sparus majo
15	52	51.5	112	1 VPX_SIVM1	P05917 simlan immu
16	52	51.5	112	1 VPX_SIVSP	P05918 simlan immu
17	51	50.5	112	1 VPX_SIVMK	P05916 simlan immu
18	51	50.5	112	1 VPX_SIVML	P11266 simlan immu
19	50	49.5	111	1 VPX_HV2KR	Q74122 human immu
20	50	49.5	112	1 VPX_HV2CA	P24110 human immu
21	50	49.5	112	1 VPX_HV2D1	P17760 human immu
22	50	49.5	112	1 VPX_HV2RO	P06939 human immu
23	49	48.5	92	1 GONI_CAVPO	O54713 cavia porce
24	48	47.5	80	1 GONI_CLAGA	P33439 ciarias gar
25	48	47.5	315	1 ALDR_PIG	P80276 sus scrofa
26	48	47.5	635	1 VP40_HSV11	P10210 herpes simp
27	47.5	47.0	114	1 GONI_TUPGB	O95336 tupiaa gils
28	47	46.5	436	1 SH6_RAT	P31388 rattus norv
29	47	46.5	2236	1 PYR1_DROME	P05990 drosophila
30	46.5	46.0	429	1 GAG_HTLIA	P03345 human t-cel
31	46.5	46.0	429	1 GAG_HTLIC	P14076 human t-cel
32	46.5	46.0	429	1 GAG_HTLIM	P14077 human t-cel
33	46.5	46.0	433	1 GAG_HTLV2	P03346 human t-cel

## ALIGNMENTS

RESULT ID	1	GONI_SHEEP	STANDARD:	PRT:	61 AA.
AC	Q28588;				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	PROCONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I))				
DE	(LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LUTILIBERIN I); GNRH-ASSOCIATED PEPTIDE I)				
DE	(FRAGMENT).				
GN	GNRH1 OR GNRH OR LHRH.				
OS	Ovis aries (Sheep).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Caprinae; Ovis.				
RN	[1]				
RP	SEQUENCE OF 12-61 FROM N.A.				
RC	STRAIN-WESTERN RANGE; TISSUE=HYPOTHALAMUS;				
RA	Rodriguez R.E., Wise M.E.;				
RL	Submitted (Oct-1993) to the EMBL/GenBank/DBD databases.				
RN	[2]				
RP	SEQUENCE OF 1-10.				
RX	MEDLINE=72094314; PubMed=4550508;				
RA	Burgus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,				
RA	Fellows R., Blackwell R., Vale W., Guillemin R.;				
RT	"Primary structure of the ovine hypothalamic luteinizing hormone-				
RT	releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass				
RT	spectrometry-decapeptide-Edman degradation).";				
RL	Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).				
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES				
CC	THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING				
CC	HORMONES.				
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL: U02517; AAA03433.1; -				
DR	PIR: A93780; RSHSG.				
DR	INTERPRO: IPR002012; -				
DR	PFAM: PF00446; GNRH; 1.				
DR	PROSITE: PS00473; GNRH; 1.				
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;				
KW	Placenta.				
FT	NON_TER	1			
FT	CHAIN	1	>61		PROCONADOLIBERIN I.
FT	PEPTIDE	1	10		GONADOLIBERIN I.
FT	PEPTIDE	14	>61		GNRH-ASSOCIATED PEPTIDE I.
FT	ACT_SITE	3	3		APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT					ACTIVITY.

34 45.5 45.0 719 1 DEND\_RAT P50617 rattus norv  
35 45 44.6 10 1 GON3\_ONCKE P20367 oncorhynchu  
36 45 44.6 74 1 GON3\_ONCKE P55246 oncorhynchu  
37 45 44.6 74 1 GON3\_ONCTS O92097 oncorhynchu  
38 45 44.6 82 1 GON3\_ONCKA P30973 oncorhynchu  
39 45 44.6 82 1 GON3\_SALSA P35629 salmo salar  
40 45 44.6 82 1 GON3\_SALTR P45653 salmo trutt  
41 45 44.6 89 1 GON3\_PORNO P51922 porichthys  
42 45 44.6 90 1 GON3\_HAPBU P45652 haplochromi  
43 45 44.6 90 1 GON3\_PAGMA P51921 pagrus majo  
44 45 44.6 90 1 GON3\_SPAUV P51923 sparus majo  
45 45 44.6 94 1 GON3\_CARAU P51917 carassius a

FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).  
FT NON\_TER 61  
SO SEQUENCE 61 AA: 6828 MW: 63962A1AE319B8F0 CRC64:

Query Match 57.4%; Score 58; DB 1; Length 61;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
1111111111  
Db 2 HWSYGLRPG 10

## RESULT 2

GN1\_MESAU STANDARD; PRT; 63 AA.  
ID GN1\_MESAU  
AC 009163;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)]  
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]  
DE (FRAGMENT).  
GN GNRH1 OR GNRH OR LHRH.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Jensen H.T., Stevens P.J., Zeitler P., Lehman M.N.;  
RU Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U01938; AAB51302.1; -.  
DR INTERPRO: IPR002012; -.  
DR PFAM: PF00446; GNRH.1.  
DR PROSITE: PS00473; GNRH.1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta.  
KM  
FT NON\_TER 1 1  
FT CHAIN 1 1 >63 PROGNADOLIBERIN I.  
FT PEPTIDE 1 1 10 GONADOLIBERIN I.  
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY SIMILARITY).  
FT ACT\_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY (BY SIMILARITY).  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).  
FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY SIMILARITY).  
FT NON\_TER 63 63  
FT SEQUENCE 63 AA: 7370 MW: FC94995676F77180 CRC64:

Query Match 57.4%; Score 58; DB 1; Length 63;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
1111111111  
Db 2 HWSYGLRPG 10

## RESULT 3

GN1\_MACMU STANDARD; PRT; 67 AA.  
ID GN1\_MACMU  
AC P55247;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)]  
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]  
DE (FRAGMENT).  
GN GNRH1 OR GNRH OR LHRH.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Cercopithecinae; Macaca.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-HYPOTHALAMUS;  
RX MEDLINE-95124501; PubMed-7545971;  
RA Ma Y.J., Costa M.E., Ojeda S.R.;  
RT "Developmental expression of the genes encoding transforming growth factor alpha and its receptor in the hypothalamus of female rhesus macaques."  
RL Neuroendocrinology 60:346-359(1994).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.  
CC -----  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: S75918; AAB33096.1; -.  
DR INTERPRO: IPR002012; -.  
DR PFAM: PF00446; GNRH.1.  
DR PROSITE: PS00473; GNRH.1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Signal.  
KM  
FT NON\_TER 1 1  
FT SIGNAL <1 5  
FT CHAIN 6 6 >67 PROGNADOLIBERIN I.  
FT PEPTIDE 6 6 15 GONADOLIBERIN I.  
FT PEPTIDE 19 >67 GNRH-ASSOCIATED PEPTIDE I.  
FT ACT\_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY (BY SIMILARITY).  
FT MOD\_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).  
FT MOD\_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY SIMILARITY).  
FT NON\_TER 67 67  
FT SEQUENCE 67 AA: 7573 MW: 505394DAA261A3F2 CRC64:

Query Match 57.4%; Score 58; DB 1; Length 67;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
1111111111  
Db 7 HWSYGLRPG 15



```

RESULT 4
GONL_XENLA STANDARD: PRT: 89 AA.
AC P45656;
DR 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FOREBRAIN;
RA MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Wray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28040; AAA9728.1; -
DR INTERPRO; IPR002012; -
DR PFAW; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 89
FT PEPTIDE 24 33
FT CHAIN 37 89
FT PEPTIDE 37 89
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 33 33
SQ SEQUENCE 89 AA; 10246 MW; 6FAF36FBAED04284 CRC64;

Query Match 57.4%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HMSGYLRPG 10
DB 25 HMSGYLRPG 33

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=87069928; PubMed=3024317;
RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA Phillips H.S., Nikolic K., Seeburg P.H.;
RT "A deletion truncating the gonadotropin-releasing hormone gene is
RT responsible for hypogonadism in the hpg mouse."
RL Science 234:1366-1371(1986).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -----
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14872; AAA37717.1; -
DR MGD; MGI:95769; GNRH.
DR INTERPRO; IPR002012; -
DR PFAW; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 21
FT CHAIN 22 90
FT PEPTIDE 22 31
FT PEPTIDE 35 90
FT ACT_SITE 24 24
FT MOD_RES 22 22
FT MOD_RES 31 31
FT MOD_RES 31 31
SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 57.4%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HMSGYLRPG 10
DB 23 HMSGYLRPG 31

RESULT 6
GONL_PIG STANDARD: PRT: 91 AA.
AC P49921;
DR 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I).
GN GNRH1 OR GNRH.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HYPOPHALAMUS;
RA Weesner G.D., Mattern R.L., Becker B.A.;
RL Submitted (MAY-1994) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 24-33.
RX MEDLINE=72114303; PubMed=4946067;
RX Baba Y., Matsuo H., Schnally A.V.;

```

```

RT      "Structure of the porcine LH- and FSH-releasing hormone. II.
RT      Confirmation of the proposed structure by conventional sequential
RT      analyses."
RL      Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN      [3]
RP      SYNTHESIS OF GONADOLIBERIN.
RX      MEDLINE=72065376; PubMed=4942726;
RA      Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT      "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RT      phase method."
RL      Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN      [4]
RP      SYNTHESIS OF GONADOLIBERIN.
RX      MEDLINE=72117544; PubMed=4946275;
RA      Baba Y., Arimura A., Schally A.V.;
RT      "On the tryptophan residue in porcine LH and FSH-releasing hormone."
RT      Biochem. Biophys. Res. Commun. 45:483-487(1971).
CC      -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC      THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC      HORMONES.
CC      -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L32864; AAA31066.1; -.
DR      PIR; A01411; RHPGG.
DR      INTERPRO: IPR002012; -.
DR      PIR: PF00446; GNRH; 1.
DR      PROSITE: PS00473; GNRH; 1.
KW      Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW      Placenta; Signal.
FT      SIGNAL 1 23
FT      CHAIN 24 91 PROGONADOLIBERIN I.
FT      PEPTIDE 24 33 GNRH-ASSOCIATED PEPTIDE I.
FT      PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
FT      ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT      ACTIVITY.
FT      MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT      MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ      SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

```

```

Query Match 57.4%; Score 58; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 HWSYGLRPG 10
DB      25 HWSYGLRPG 33

RESULT 7
GONL_HUMAN STANDARD; PRT; 92 AA.
AC      F01148;
DT      21-JUL-1986 (Rel. 01, Created)
DT      01-APR-1988 (Rel. 07, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE      (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE      HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED
DE      PEPTIDE I].
GN      GNRH1 OR GNRH OR LHRH.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN      [1]

```

```

RP      SEQUENCE FROM N.A.
RX      MEDLINE=8936682; PubMed=2671939;
RA      Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT      "The complete nucleotide sequence of the human gonadotropin-releasing
RT      hormone gene."
RL      Nucleic Acids Res. 17:6403-6403(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86094338; PubMed=2867548;
RA      Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT      "Isolation of the gene and hypothalamic cDNA for the common precursor
RT      of gonadotropin-releasing hormone and prolactin release-inhibiting
RT      factor in human and rat."
RL      Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=85012739; PubMed=6090951;
RA      Seeburg P.H., Adelman J.P.;
RT      "Characterization of cDNA for precursor of human luteinizing hormone
RT      releasing hormone."
RL      Nature 311:666-668(1984).
RN      [4]
RP      SEQUENCE OF 24-33.
RX      MEDLINE=83126573; PubMed=6760865;
RA      Tan L., Rousseau P.;
RT      "The chemical identity of the immunoreactive LHRH-like peptide
RT      biosynthesized in the human placenta."
RL      Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC      -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC      THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC      HORMONES.
CC      -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS),
CC      LUTREPUSE OR LUTRELEF (FERRING PHARMACEUTICALS) AND RELISORM
CC      (SERONO).
CC      -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X01059; CAA25526.1; -.
DR      EMBL; M12578; AAA35916.1; -.
DR      EMBL; X15215; CAA33285.1; -.
DR      PIR; A01410; RHHUG.
DR      PIR; A26173; A26173.
DR      PIR; S05308; S05308.
DR      PIR; 152760; -.
DR      INTERPRO: IPR002012; -.
DR      PIR: PF00446; GNRH; 1.
DR      PROSITE: PS00473; GNRH; 1.
KW      Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW      Placenta; Pharmaceutical; Signal.
FT      SIGNAL 1 23
FT      CHAIN 24 92 PROGONADOLIBERIN I.
FT      PEPTIDE 24 92 GNRH-ASSOCIATED PEPTIDE I.
FT      PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT      ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT      ACTIVITY.
FT      MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT      MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT      CONFLICT 16 16 W -> S (IN REF. 3).
SQ      SEQUENCE 92 AA; 10380 MW; 30A7221B076FA79 CRC64;

```

```

Query Match 57.4%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2 HWSYGLRPG 10

```

Db 25 HWSYGLRPG 33

```

RESULT 8
GONL_RAT STANDARD: PRT: 92 AA.
ID GONL_RAT
AC P07490;
DT 01-APR-1988 (rel. 07, Last sequence update)
DT 01-APR-1988 (rel. 07, Last sequence update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
DE I].
GN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN
RP SEQUENCE FROM N.A.
RX MEDLINE-86094338; PubMed-2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-89384661; PubMed-2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
RT hypothalamic expression.";
RL Mol. Endocrinol. 3:1257-1262(1989).
RN
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE-THYMUS;
RA MEDLINE-93105480; PubMed-1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RX TISSUE-HEART;
RA MEDLINE-87149087; PubMed-3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus.";
RL Science 235:1514-1517(1987).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S50879; AAB24572.1; -
CC EMBL: M12570; AAA41263.1; -
CC EMBL: M1670; AAA41264.1; -
CC EMBL: M15527; AAA42141.1; ALT-SEQ.
CC EMBL: M15529; AAA42139.1; -
CC EMBL: M15528; -; NOT_ANNOTATED_CDS.
CC PIR: B26173; RHRTG.
CC DR PIR: A48410; A48410.
CC DR INTERPRO: IPR002012; -

```

```

DR PFAM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DABA3EB3 CRC64;
Query Match 57.4%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
RESULT 9
GONL_TUPGB STANDARD: PRT: 92 AA.
ID GONL_TUPGB
AC Q95335;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRH1 OR GNRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupalidae; Tupala.
RN
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-HYPOTHALAMUS;
RA MEDLINE-97079639; PubMed-8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.;
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U63326; AAB16837.1; -
CC DR INTERPRO: IPR002012; -
CC DR PFAM: PF00446; GNRH: 1.
CC DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT ACT_SITE 26 26
BY SIMILARITY.
PROGNADOLIBERIN I.
GONADOLIBERIN I.
GNRH-ASSOCIATED PEPTIDE I.
APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.

```

```

FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SEQUENCE 92 AA; 10197 MW; 4FDB2C58CF5F63B CRC64;
SQ
Query Match 57.4%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
RESULT 10
GONI_ALLMI STANDARD; PRT; 10 AA.
ID GONI_ALLMI P20407;
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I)
DE (LULIBERIN I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorineae; Alligator.
RN [1]
RP SEQUENCE.
RC TISSUE-BRAIN;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., Moroy J.E., Park M.,
lance V., Swanson P., Rivier J.E., Sherwood N.M.,
"Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RT Regul. Pept. 33:105-116(1991).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: A60065; RHA01.
DR INTERPRO: IPR002012; -.
DR PIR: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;
Query Match 53.5%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.064;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10
RESULT 11
GONI_CHICK STANDARD; PRT; 92 AA.
ID GONI_CHICK P20407;
AC P37042; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEGHORN;
RX MEDLINE=94059355; PubMed=7902095;
RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
RT "Characterization of the chicken preprogonadotropin-releasing
RT hormone-I gene.";
RN J. Mol. Endocrinol. 11:19-29(1993).
RN [2]
RP SEQUENCE OF 24-33.
RC TISSUE-HYPOTHALAMUS;
RX MEDLINE=82265778; PubMed=7050119;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RT hormone. II. Isolation and characterization.";
RN J. Biol. Chem. 257:10729-10732(1982).
RN [3]
RP SEQUENCE OF 24-33.
RC TISSUE-HYPOTHALAMUS;
RA King J.A., Millar R.P.;
RT "Structure of avian hypothalamic gonadotropin-releasing hormone.";
RN S. Afr. J. Sci. 78:124-125(1982).
RN [4]
RP SYNTHESIS OF 24-33.
RX MEDLINE=82265777; PubMed=7050119;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RT hormone. I. Structural determination on partially purified
RT material.";
RN J. Biol. Chem. 257:10722-10728(1982).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69491; CAA49246.1; -.
DR PIR: S33507; S33507.
DR INTERPRO: IPR002012; -.
DR PIR: PF00446; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23 PROGONADOLIBERIN I.
FT CHAIN 24 92 GONADOLIBERIN I.
FT PEPTIDE 24 33 GNRH-ASSOCIATED PEPTIDE I.
FT PEPTIDE 37 92 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 24 24 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10206 MW; 61AEB7EBAF508B6A CRC64;
Query Match 53.5%; Score 54; DB 1; Length 92;
Best Local Similarity 88.9%; Pred. No. 0.055;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
RESULT 12
GONI_HAPBU STANDARD; PRT; 94 AA.
ID GONI_HAPBU P51918;
AC P51918; 093387;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)

```

```

DN      (LH-RH I.) (LULIBERIN I.).
GN      GnRH1.
OS      Haplochromis burtoni.
OC      Euparietia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC      Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC      Clariidae; Astrotellapia.
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95396797; Pubmed=7667296;
RA      White S.A., Kassten T.L., Bond C.T., Adelman J.P., Fernald R.D.;
RT      "Three gonadotropin-releasing hormone genes in one organism suggest
RT      novel roles for an ancient peptide.";
RL      Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99061842; Pubmed=9843638;
RA      White R.B., Fernald R.D.;
RT      "Ontogeny of gonadotropin-releasing hormone (GnRH) gene expression
RT      reveals a distinct origin for GnRH-containing neurons in the
RT      midbrain.";
RL      Gen. Comp. Endocrinol. 112:322-329(1998).
RN      [3]
RP      SEQUENCE OF 23-32.
RC      TISSUE=PIUTINARY;
RX      MEDLINE=95372591; Pubmed=7644702;
RA      Powell J.E.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
RA      White S.A., Francis R.C., Fernald R.D., Licht P., Warby C.,
RA      Sherwood N.M.;
RT      "Primary structure of solitary form of gonadotropin-releasing hormone
RT      (GnRH) in cichlid pituitary; three forms of GnRH in brain of cichlid
RT      and pumpkinseed fish.";
RL      Regul. Pept. 57:43-53(1995).
CC      -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
CC      RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-
CC      GONADAL AXIS.
CC      -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
CC      TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOPHYSAL AXONS.
CC      -1- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.
CC      -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL; U31865; AAC59691.1; -.
DR      EMBL; AF076961; AAC27716.1; -.
DR      INTERPRO: IPR002012; -.
DR      PFMW: PF00446; GnRH; 1.
DR      PROSITE: PS00473; GnRH; 1.
KW      Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW      Signal; Multigene family.
FT      SIGNAL 1
FT      CHAIN 23
FT      PEPTIDE 23 32
FT      PEPTIDE 36 94
FT      MOD_RES 23 23
FT      MOD_RES 32 32
FT      CONFLICT 86 94
SQ      SEQUENCE 94 AA; 10382 MM; E57DBA8333278D7 CR664;

```

RESULT	13
GONL_PAGMA	
ID	GONL_PAGMA
AC	P70074;
DT	15-JUL-1998 (Rel. 36, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	15-DEC-1998 (Rel. 37, Last annotation update)
DE	GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE	(LH-RH I) (LILIBERIN I).
OS	Pagrus major (red sea bream) (Chrysophrys major).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
CC	Sparidae; Chrysophrys.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSE-BRAIN:
RA	Oukuzawa K., Granneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
RL	Submitted (SEP-1996) to the EMBL/Genebank/DBJ databases.
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPHINS (BY
CC	SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; D86582; BAAL3129.1; -
DR	INTERPRO; IPRO02012; -
DR	PFAM; PF00446; GNRH; 1.
DR	PROSITE; PS00473; GNRH; 1.
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
FT	Signal; Multigene family.
FT	SIGNAL 1 23
FT	CHAIN 1 24 33
FT	PEPTIDE 24 33
FT	PERIODONE CARBOXYLIC ACID (BY
FT	MOD_RES 37 95
FT	MOD_RES 24 24
FT	MOD_RES 33 33
FT	MOD_RES 33 33
FT	MOD_RES 33 33
SEQ	SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;
Query Match	51.5%; Score 52; DB 1; Length 95;
Best Local Similarity	88.9%; Pred. NO. 1;
Matches	8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	2 HMSYGLRPG 10
Dd	25 HWSTGLSPG 33
RESULT	14
GONL_SPAU	
ID	GONL_SPAU
AC	P51918;
DT	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	15-DEC-1998 (Rel. 37, Last annotation update)
DE	GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE	(LH-RH I) (LILIBERIN I) (SBGNH).
OS	Sparus aurata (Gilthead sea bream).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
CC	Sparidae; Sparus.

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RX MEDLINE=95268499; PubMed=7749463;
RA Gotthelf Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-
  releasing hormone from the gillhead seabream (Sparus aurata).";
RL Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE-BRAIN:
RX MEDLINE=95083645; PubMed=7991588;
RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
  brains of one species.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U30320; AAA75469.1; -.
DR INTERPRO: IPR002012; -.
DR PFM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 25
FT CHAIN 26 95 PROGONADOLIBERIN I.
FT PEPTIDE 26 95 GONADOLIBERIN I.
FT MOD_RES 39 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 35 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
SQ SEQUENCE 95 AA; 10753 MW; 49313FD6FDB87DA CRC64;

Query Match
Best Local Similarity 51.5%; Score 52; DB 1; Length 95;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
   11111111
Db 27 HWSYGLSPG 35

RESULT 15
VPX_SIVM1 STANDARD; PRT; 112 AA.
AC P05917;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE VPX PROTEIN (X ORF PROTEIN).
GN VPX.
OS Simian immunodeficiency virus (Mm142-83 isolate) (SIV-MAC).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87287230; PubMed=3649576;
RA Chakrabarti L., Guyader M., Allison M., Daniel M.D., Desrosiers R.C.,
  Tiollais P., Sonigo P.;
RT "Sequence of simian immunodeficiency virus from macaque and its
  relationship to other human and simian retroviruses.";
RL Nature 328:543-547(1987).
CC -1- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00277; CAA68382.1; -.
DR PIR: D28887; ASLJX3.
DR HIV: M16403; VPXMM142.
DR INTERPRO: IPR000012; -.
DR PFM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVVPRVPX.
KW AIDS.
SQ SEQUENCE 112 AA; 12906 MW; 46F564F45AFED960 CRC64;

Query Match
Best Local Similarity 51.5%; Score 52; DB 1; Length 112;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRPGRP 15
   11111111
Db 94 HGAAGWRPGPPPP 107

Search completed: March 2, 2001, 11:01:20
Job time: 435 sec
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:55:40 ; Search time 56.87 Seconds  
(without alignments)  
19.103 Million cell updates/sec

Title: us-09-306-689-5

Perfect score: 101

Sequence: 1 XHMSYGLRPGRRPPPC 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	57.4	10	1 RHPGC	gonadoliberin - pi
2	58	57.4	10	1 RHPGC	gonadoliberin - sh
3	58	57.4	10	1 RHPGC	gonadoliberin prec
4	58	57.4	89	2 I51423	gonadoliberin prec
5	58	57.4	90	1 RHMSG	gonadoliberin prec
6	58	57.4	92	1 RHMSG	gonadoliberin prec
7	58	57.4	92	1 RHMSG	gonadoliberin prec
8	55	54.5	969	2 T15446	gonadoliberin I -
9	54	53.5	10	1 RHAQ1	gonadoliberin I -
10	54	53.5	92	2 I50644	gonadoliberin I pr
11	52	51.5	98	2 I50739	gonadotropin-relea
12	52	51.5	106	2 A30996	gonadotropin-relea
13	52	51.5	112	1 ASLJX3	gonadotropin-relea
14	52	51.5	112	1 ASLJX3	gonadotropin-relea
15	52	51.5	112	1 ASLJX3	gonadotropin-relea
16	51	50.5	112	1 ASLJX2	gonadotropin-relea
17	50	49.5	112	1 ASLJX2	gonadotropin-relea
18	50	49.5	112	2 SI2155	gonadotropin-relea
19	50	49.5	112	2 SI2155	gonadotropin-relea
20	50	49.5	112	2 SI2155	gonadotropin-relea
21	50	49.5	112	2 SI2155	gonadotropin-relea
22	49	48.5	373	2 A47234	gonadotropin-relea
23	49	48.5	2129	2 T27431	gonadotropin-relea
24	48	47.5	80	1 RHIDS	gonadotropin-relea
25	48	47.5	316	2 A59021	gonadotropin-relea
26	48	47.5	598	2 T48822	gonadotropin-relea
27	48	47.5	629	2 S31174	gonadotropin-relea
28	48	47.5	635	1 NM5EW6	gonadotropin-relea
29	47	46.5	436	2 JN0591	gonadotropin-relea

30	47	46.5	587	2 T19893	hypothetical prote
31	47	46.5	587	2 T29324	hypothetical prote
32	47	46.5	2236	1 QZFF	hypothetical prote
33	46.5	46.0	429	1 FOLJCH	probable gag poly
34	46.5	46.0	429	1 FOLJCN	gag polypeptide -
35	46.5	46.0	429	1 S06073	gag polypeptide -
36	46.5	46.0	433	1 FOLJH2	gag polypeptide -
37	46	45.5	477	2 F64805	RhA protein (lipo
38	46	45.5	972	2 T47922	anthranilate phosph
39	45.5	45.0	112	2 S19976	Ig kappa chain V r
40	45.5	45.0	613	2 T42671	hypothetical prote
41	45	44.6	10	2 A21114	gonadotropin relea
42	45	44.6	74	2 I51092	gonadotropin relea
43	45	44.6	82	2 I51180	gonadotropin relea
44	45	44.6	82	2 I51355	gonadotropin relea
45	45	44.6	82	2 I51365	gonadotropin relea

#### ALIGNMENTS

RESULT 1  
RHPGC  
gonadoliberin - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 18-Mar-1997  
C:Accession: A01411  
R:Baba, Y.; Matsuo, H.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 44, 459-463, 1971  
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of  
A:Reference number: A90172; MUID:72114303  
A:Accession: A01411  
A:Molecule type: protein  
A:Residues: 1-10 <BAB>  
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 45, 822-827, 1971  
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me  
A:Reference number: A90176; MUID:72065376  
A:Contents: annotation; synthesis  
A:Note: The synthetic and natural hormones have the same physicochemical and biologic  
R:Baba, Y.; Arimura, A.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 45, 483-487, 1971  
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.  
A:Reference number: A90175; MUID:72117544  
A:Contents: annotation  
A:Note: Trp-3 appears to be essential for biological activity  
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F1/Modified site: pyroglutamic carboxylic acid (Gln) #status experimental  
F10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 57.4%; Score 58; DB 1; Length 10;  
Best local similarity 100.0%; Pred. No. 0.038;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSYGLRPG 10  
DB 2 HMSYGLRPG 10

RESULT 2  
RHPGC  
gonadoliberin - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Mar-1997  
C:Accession: A93780; A01411  
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;  
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972  
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa  
A:Reference number: A93780; MUID:72094314  
A:Accession: A93780

A:Molecule type: protein  
 A:Residues: 1-10 <B0R>  
 A>Note: the natural and synthetic hormones have the same biological activity  
 C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo  
 C:Superfamily: gonadoliberin  
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
 F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
 F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 57.4%; Score 58; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.038;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||||  
 DB 2 HWSYGLRPG 10

RESULT 3  
 178541  
 gonadoliberin precursor - rhesus macaque (fragment)  
 N:Alternate names: luteinizing hormone releasing hormone  
 C:Species: Macaca mulatta (rhesus macaque)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
 C:Accession: I78541  
 R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.  
 Neuroendocrinology 60, 346-359, 1994  
 A:Title: Developmental expression of the genes encoding transforming growth factor alpha  
 A:Reference number: I58134; MUID:95124501  
 A:Status: preliminary;  
 A:Molecule type: mRNA  
 A:Residues: 1-67 <RES>  
 A:Cross-references: GB:S75918; NID:g912831; PIDN:AAA33096.1; PID:g912832  
 C:Superfamily: gonadoliberin

Query Match 57.4%; Score 58; DB 2; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||||  
 DB 7 HWSYGLRPG 15

RESULT 4  
 151423  
 gonadoliberin precursor - African clawed frog  
 N:Alternate names: luteinizing hormone releasing hormone  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
 C:Accession: I51423  
 R:Hayes, W.P.; Wray, S.; Battey, J.F.  
 Endocrinology 134, 1835-1845, 1994  
 A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved doma  
 A:Reference number: I51423; MUID:94185563  
 A:Accession: I51423  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-89 <HAY>  
 A:Cross-references: GB:L28040; NID:g496291; PIDN:AAA9728.1; PID:g496292  
 C:Genetics:  
 A:Gene: GnRH-I  
 C:Superfamily: gonadoliberin

Query Match 57.4%; Score 58; DB 2; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 0.34;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10

DB 25 HWSYGLRPG 33  
 |||||||

RESULT 5  
 RHMSG  
 gonadoliberin precursor - mouse  
 N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasi  
 N:Contains: gonadoliberin; gonadoliberin-associated protein (GAP)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Dec-1993 #sequence\_revision 18-Mar-1997 #text\_change 18-Jun-1999  
 C:Accession: A47578  
 R:Wason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli  
 Science 234, 1366-1371, 1986  
 A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible  
 A:Reference number: A47578; MUID:87069928  
 A:Accession: A47578  
 A:Molecule type: DNA  
 A:Residues: 1-90 <MAS>  
 A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175  
 C:Genetics:  
 A:Introns: 45/3; 77/3  
 C:Function:  
 A:Description: gonadoliberin stimulates pituitary secretion of luteinizing hormone and follicle  
 A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activi  
 C:Superfamily: gonadoliberin  
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:22-31/Product: gonadoliberin #status predicted <GLP>  
 F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>  
 F:22/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predic  
 F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 57.4%; Score 58; DB 1; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 0.34;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||||  
 DB 23 HWSYGLRPG 31

RESULT 6  
 RHMSG  
 gonadoliberin precursor - human  
 N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasi  
 N:Contains: gonadoliberin-associated protein (GAP); progadoliberin  
 C:Species: Homo sapiens (man)  
 C:Date: 17-Mar-1987 #sequence\_revision 21-Jul-1995 #text\_change 18-Jul-1999  
 C:Accession: S05308; A26173; A93342; A90108; A01410; S45718  
 R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.  
 Nucleic Acids Res. 17, 6403-6404, 1989  
 A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone  
 A:Reference number: S05308; MUID:89366682  
 A:Accession: S05308  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-92 <HAY>  
 A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g332956  
 R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
 A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gona  
 A:Reference number: A94090; MUID:86094338  
 A:Accession: A26173  
 A:Molecule type: mRNA  
 A:Residues: 1-92 <ADE>  
 A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749  
 R:Seeburg, P.H.; Adelman, J.P.  
 Nature 311, 666-668, 1984  
 A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasin  
 A:Reference number: A93342; MUID:85012739



A:Accession: A93342  
A:Molecule type: mRNA  
A:Residues: 1-15, 'S', 17-92 <SEE>  
A:Cross-references: GB:X01059; NID:934356; PIDN:CAA2526.1; PID:934357  
A:Experimental source: placenta  
R:Tan, L.; Rousseau, P.  
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982  
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in  
A:Reference number: A90108; MUID:83126573  
A:Accession: A90108  
A:Molecule type: Protein  
A:Residues: 24-33 <TAN>  
A:Experimental source: placental trophoblasts  
R:Leibovitz, D.; Koch, Y.; Piltzer, F.; Fridkin, M.; Dantes, A.; Baumbaister, W.; Amsterda  
FBS Lett. 346, 203-206, 1994  
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th  
A:Reference number: S45718; MUID:94283597  
A:Contents: annotation; degradation pathway of synthetic hormone  
C:Genetics:  
A:Gene: GDB:GNRH; LHRH; GRH  
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760  
A:Map position: 8p21-8p11.2  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyrogutamic acid  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-92/Product: progadoliberin #status predicted <PGN>  
F:24-33/Product: gonadoliberin #status experimental <MAT>  
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 57.4%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 HMSGYLRPG 10  
DB 25 HMSGYLRPG 33  
|||||  
|

RESULT 7  
RHRGTG  
gonadoliberin precursor - rat  
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormo  
N:Contains: gonadoliberin; prolactin release-inhibiting factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1999  
C:Accession: A40147; B26173; A48410  
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.  
MOL. Endocrinol. 3, 1257-1262, 1989  
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex  
A:Reference number: A40147; MUID:89384661  
A:Accession: A40147  
A:Molecule type: DNA  
A:Residues: 1-92 <BON>  
A:Cross-references: GB:M31670; NID:9204447; PIDN:AAA1264.1; PID:9204448  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot  
A:Reference number: A9090; MUID:86094338  
A:Accession: B26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12579; NID:9204445; PIDN:AAA1263.1; PID:9204446  
R:Walter, C.C.; Marchetti, B.; LeBoeuf, R.D.; Bialock, J.E.  
Cell. Mol. Neurobiol. 12, 447-454, 1992  
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone  
A:Reference number: A48410; MUID:93105480

A:Accession: A48410  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-92 <MAT>  
A:Cross-references: GB:S50870; NID:9262059; PIDN:AAB24572.1; PID:9262060  
A:Experimental source: thymus  
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIF:121083)  
C:Genetics:  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: stimulates pituitary secretion of luteotropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activi  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyrogutamic acid  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-92/Product: progadoliberin #status predicted <PGN>  
F:24-33/Product: gonadoliberin #status predicted <Gln>  
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 57.4%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 HMSGYLRPG 10  
DB 25 HMSGYLRPG 33  
|||||  
|

RESULT 8  
T15446  
hypothetical protein C07G1.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T15446  
R:Hawkins, J.  
Submitted to the EMBL Data Library, May 1996  
A:Description: The sequence of C. elegans cosmid C07G1.  
C:Accession number: Z18352  
A:Accession: T15446  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-969 <HAM>  
A:Cross-references: EMBL:U58751; NID:91326379; PID:91326381; PIDN:AAB00657.1; GSPDB:G  
A:Experimental source: strain Bristol N2; clone C07G1  
C:Genetics:  
A:Gene: CESP:C07G1.4  
A:Map position: 4  
A:Introns: 110/3; 149/2; 515/1; 564/1; 600/2; 690/2; 755/3; 843/3; 884/3; 940/2

Query Match 54.5%; Score 55; DB 2; Length 969;  
Best Local Similarity 75.0%; Pred. No. 9.2;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
OY 4 SYGLRPPRPP 15  
DB 727 SHGLAPRPP 738  
|:|:| |  
|

RESULT 9  
RHAQI  
gonadoliberin I - American alligator  
N:Alternate names: gonadotropin-releasing hormone I  
C:Species: Alligator mississippiensis (American alligator)  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Mar-1997  
C:Accession: A60066  
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan  
Regul. Pept. 33, 105-116, 1991  
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains  
A:Reference number: A60066; MUID:91352338

A:Accession: A60066  
A:Molecule type: protein  
A:Residues: 1-10 <LOV>  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 53.5%; Score 54; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.13;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
| | | | | | | |  
Db 2 HWSYGLQPG 10

RESULT 10  
I50644

gonadoliberin I precursor - chicken

N:Alternate names: gonadotropin-releasing hormone I

C:Species: Gallus gallus (chicken)

C>Date: 21-Feb-1997 #sequence\_rev150644; S33507

C:Accession: I50644; S33507

R:Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.

J. Mol. Endocrinol. 11, 19-29, 1993

A>Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene.

A:Reference number: I50644; PMID:94059355

A:Accession: I50644

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-92 <DD2>

A:Cross-references: EMBL:X69491; NID:9496326; PIDN:CAA49246.1; PID:g311612

C:Genetics: 47/3; 79/3

A:Insertions: 79/3

C:Superfamily: gonadoliberin

Query Match 53.5%; Score 54; DB 2; Length 92;  
Best Local Similarity 88.9%; Pred. No. 1.2;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
| | | | | | | |  
Db 25 HWSYGLQPG 33

RESULT 11  
I50739

gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)

C:Species: Haplochromis burtoni

C>Date: 13-Sep-1996 #sequence\_rev150739; S33507

C:Accession: I50739

R:White, S.A.; Kastan, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.

Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995

A>Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles

A:Reference number: I50739; PMID:9536797

A:Accession: I50739

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-98 <WHI>

A:Cross-references: EMBL:U11865; NID:9905398; PIDN:AAC59691.1; PID:g905399

C:Superfamily: gonadoliberin

Query Match 51.5%; Score 52; DB 2; Length 98;  
Best Local Similarity 88.9%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
| | | | | | | |  
Db 24 HWSYGLSPG 32

RESULT 12

A30996

orf-X protein - simian immunodeficiency virus (isolate Mne) (fragments)

N:Alternate names: vpx protein

C:Species: simian immunodeficiency virus, SIV

C>Date: 07-Sep-1990 #sequence\_rev150996; S33507

C:Accession: A30996

R:Henderson, L.E.; Sowder, R.C.; Copeland, T.D.; Benveniste, R.E.; Oroszian, S.

Science 241, 199-201, 1988

A>Title: Isolation and characterization of a novel protein (X-ORF product) from SIV a

A:Reference number: A30996; PMID:88264422

A:Accession: A30996

A:Molecule type: protein

A:Residues: 1-106 <HEN>

C:Genetics:

A:Gene: vpx

C:Superfamily: AIDS vpu protein

Query Match 51.5%; Score 52; DB 2; Length 106;  
Best Local Similarity 64.3%; Pred. No. 2.5;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRPPP 15  
| : | | | | | |  
Db 88 HGAGGWRPGRPPP 101

RESULT 13  
ASLJX3

vpu protein - simian immunodeficiency virus (macaque isolate)

N:Alternate names: orf-X protein

C:Species: simian immunodeficiency virus, SIV

C>Date: 30-Jun-1989 #sequence\_rev150996; S33507

C:Accession: D28887

R:Chakrabarti, L.; Guyader, M.; Alizon, M.; Daniel, M.D.; Desrosiers, R.C.; Tiollais, N.

Nature 328, 543-547, 1987

A>Title: Sequence of simian immunodeficiency virus from macaque and its relationship

A:Reference number: A28887; PMID:87287230

A:Accession: D28887

A:Molecule type: DNA

A:Residues: 1-112 <CHA>

A:Cross-references: GB:Y00277; GB:M16403; NID:961730; PIDN:CAA68382.1; PID:g61734

C:Genetics:

A:Gene: vpu

C:Superfamily: AIDS vpu protein

Query Match 51.5%; Score 52; DB 1; Length 112;  
Best Local Similarity 64.3%; Pred. No. 2.6;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRPPP 15  
| : | | | | | |  
Db 94 HGAGGWRPGRPPP 107

RESULT 14  
S03066

gene X protein - human T-cell lymphotropic virus type 4

C:Species: human T-cell lymphotropic virus type 4, HTLV-4

C>Date: 21-Nov-1993 #sequence\_rev150306; S33507

C:Accession: S03066

R:Hahn, B.H.; Kong, L.I.; Lee, S.W.; Kumar, P.; Taylor, M.E.; Arya, S.K.; Shaw, G.M.

Nature 300, 184-186, 1987

A>Title: Relation of HTLV-4 to simian and human immunodeficiency-associated viruses.

A:Reference number: S03066

A:Accession: S03066

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-112 <HAH>  
 A:Cross-references: EMBL:X06391; NID:g61580; PIDN:CAA29688.1; PID:g61581  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1987  
 C:Superfamily: AIDS vpu protein

Query Match 51.5%; Score 52; DB 2; Length 112;  
 Best Local Similarity 64.3%; Pred. No. 2.6;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPGRRPP 15  
 I : | | | | |  
 Db 94 HGAGGWRPGRRPP 107

RESULT 15  
 ASLUST  
 vpu protein - simian immunodeficiency virus SIVagm (type 3, isolate STRV-3agm)  
 N:Alternate names: orf-X protein  
 C:Species: simian immunodeficiency virus SIVagm  
 C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 31-Jan-1997  
 C:Accession: D26737  
 R:Hirsch, V.  
 submitted to GenBank, June 1987  
 A:Reference number: A26739  
 A:Accession: D26737  
 A:Molecule type: DNA  
 A:Residues: 1-112 <HIR>  
 A:Genetics:  
 A:Gene: vpu  
 C:Superfamily: AIDS vpu protein  
 C:Keywords: AIDS; immunodeficiency

Query Match 50.5%; Score 51; DB 1; Length 112;  
 Best Local Similarity 64.3%; Pred. No. 3.5;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRPGRRPP 15  
 I : | | | | |  
 Db 94 HRAGGWRPGRRPP 107

Search completed: March 2, 2001, 10:55:41  
 Job time: 261 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:36 ; Search time 47.48 Seconds  
(without alignments)  
6.429 Million cell updates/sec

Title: US-09-306-689-4

Sequence: 1 XHMSYGLRPGSSPPPC 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents,AA:\*  
1: /cgn2\_6/prodata/2/1aa/3A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/PCUTUS\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	99.0	17	1	US-08-188-223-6 Sequence 6, Appl1
2	103	99.0	17	3	US-08-968-466-6 Sequence 6, Appl1
3	83.5	80.3	16	1	US-08-188-223-2 Sequence 2, Appl1
4	83.5	80.3	16	3	US-08-968-466-2 Sequence 2, Appl1
5	65	62.5	52	3	US-08-458-814-6 Sequence 6, Appl1
6	65	62.5	55	3	US-08-458-814-7 Sequence 7, Appl1
7	62	59.6	14	1	US-08-453-588-23 Sequence 23, Appl1
8	62	59.6	14	3	US-08-521-079-23 Sequence 23, Appl1
9	62	59.6	17	1	US-07-690-983D-18 Sequence 18, Appl1
10	62	59.6	18	1	US-07-690-983D-18 Sequence 28, Appl1
11	62	59.6	49	1	US-08-387-156-4 Sequence 4, Appl1
12	62	59.6	49	2	US-08-694-865-4 Sequence 4, Appl1
13	62	59.6	49	3	US-08-878-748-4 Sequence 4, Appl1
14	62	59.6	49	3	US-09-124-491-4 Sequence 4, Appl1
15	62	59.6	544	1	US-08-387-156-10 Sequence 10, Appl1
16	62	59.6	544	2	US-08-694-865-10 Sequence 10, Appl1
17	62	59.6	544	2	US-08-878-748-10 Sequence 10, Appl1
18	62	59.6	544	2	US-08-878-748-10 Sequence 10, Appl1
19	62	59.6	544	2	US-09-124-491-10 Sequence 10, Appl1
20	62	59.6	699	3	US-08-694-865-16 Sequence 16, Appl1
21	62	59.6	699	3	US-09-124-491-16 Sequence 16, Appl1
22	62	59.6	977	1	US-08-387-156-8 Sequence 8, Appl1
23	62	59.6	977	2	US-08-694-865-8 Sequence 8, Appl1
24	62	59.6	977	2	US-08-878-748-8 Sequence 8, Appl1
25	59	56.7	12	1	US-09-124-491-8 Sequence 8, Appl1
26	59	56.7	12	1	US-08-453-588-24 Sequence 24, Appl1
27	59	56.7	14	1	US-08-521-079-24 Sequence 24, Appl1
28	59	56.7	14	1	US-07-690-983D-22 Sequence 22, Appl1

29	59	56.7	14	1	US-07-690-983D-26 Sequence 26, Appl1
30	59	56.7	24	1	US-07-690-983D-43 Sequence 43, Appl1
31	59	56.7	44	1	US-07-690-983D-45 Sequence 45, Appl1
32	59	56.7	84	1	US-07-690-983D-47 Sequence 47, Appl1
33	58	55.8	10	1	US-07-714-540-9 Sequence 9, Appl1
34	58	55.8	10	1	US-07-690-983D-2 Sequence 2, Appl1
35	58	55.8	10	1	US-07-690-983D-32 Sequence 32, Appl1
36	58	55.8	10	1	US-08-103-022-1 Sequence 1, Appl1
37	58	55.8	10	1	US-08-184-935-6 Sequence 6, Appl1
38	58	55.8	10	1	US-08-343-883-1 Sequence 1, Appl1
39	58	55.8	10	1	US-08-000-931-5 Sequence 5, Appl1
40	58	55.8	10	1	US-08-428-488-22 Sequence 22, Appl1
41	58	55.8	10	1	US-08-341-219-11 Sequence 11, Appl1
42	58	55.8	10	1	US-08-453-588-2 Sequence 2, Appl1
43	58	55.8	10	1	US-08-453-588-4 Sequence 4, Appl1
44	58	55.8	10	1	US-08-453-588-6 Sequence 6, Appl1
45	58	55.8	10	5	5492893-1 Patent No. 5492893

## ALIGNMENTS

RESULT 1  
US-08-188-223-6  
Sequence 6, Application US/08188223  
Patent No. 5688506  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Sciablenki, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,223  
FILING DATE: 27-JAN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label=pglu  
OTHER INFORMATION: /note="pyroglutamic acid (5-oxoproline)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note="immunomimic"  
FEATURE:  
NAME/KEY: Region

LOCATION: 11..17  
OTHER INFORMATION: /note= "spacer"  
US-08-188-223-6

Query Match 99.0%; Score 103; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPGSSPPPC 17

## RESULT 2

US-08-968-466-6  
Sequence 6, Application US/08968466  
Patent No. 6132720  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Scibienski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968.466  
FILING DATE: 27-JAN-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label= pglu  
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11..17  
OTHER INFORMATION: /note= "spacer"  
US-08-968-466-6

Query Match 99.0%; Score 103; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPGSSPPPC 17

## RESULT 3

US-08-188-223-2  
Sequence 2, Application US/08188223  
Patent No. 568506  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Scibienski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188.223  
FILING DATE: 27-JAN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11..16  
OTHER INFORMATION: /note= "spacer"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label= pglu  
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"  
US-08-188-223-2

Query Match 80.3%; Score 83.5; DB 1; Length 16;  
Best Local Similarity 87.5%; Pred. No. 5.3e-05;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPG-RRPPC 16

RESULT 4  
US-08-968-466-2

```
Sequence 2, Application US/08968466
Patent No. 6132720
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Divas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,466
FILING DATE: 27-JAN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Divas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-354-8113
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..16
OTHER INFORMATION: /note= "spacer"
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= "pglu"
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoprolinone)"
US-08-968-466-2

Query Match      80.3%; Score 83.5; DB 3; Length 16;
Best Local Similarity 87.5%; Pred. No. 5.3e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY      2 HWSYGLRPGSSPPPC 17
Db      2 HWSYGLRPG-RPPPC 16

RESULT 5
US-08-458-814-6
Sequence 6, Application US/08458814
Patent No. 6103243
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J
APPLICANT: DE AIZBURDA, Henry J
APPLICANT: HOWE, Peter
APPLICANT: RAND, Keith N
TITLE OF INVENTION: ORAL VACCINES
```

```
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,814
FILING DATE: 02-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,822
FILING DATE: 18-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU86/00135
FILING DATE: 14-MAY-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH3104
FILING DATE: 25-OCT-1985
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH0566
FILING DATE: 15-MAY-1985
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/155/BIAU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-814-6

Query Match      62.5%; Score 65; DB 3; Length 52;
Best Local Similarity 83.3%; Pred. No. 0.037;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      2 HWSYGLRPGSSP 13
Db      37 HWSYGLRPGDP 48

RESULT 6
US-08-458-814-7
Sequence 7, Application US/08458814
Patent No. 6103243
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J
APPLICANT: DE AIZBURDA, Henry J
APPLICANT: HOWE, Peter
APPLICANT: RAND, Keith N
TITLE OF INVENTION: ORAL VACCINES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,814  
FILING DATE: 02-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/327,822  
FILING DATE: 18-OCT-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU86/00135  
FILING DATE: 14-MAY-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PH3104  
FILING DATE: 25-OCT-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PH0566  
FILING DATE: 15-MAY-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 60042/155/BIAN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-458-814-7

Query Match 62.5%; Score 65; DB 3; Length 55;  
Best Local Similarity 83.3%; Pred. No. 0.039;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSP 13  
DB 39 HWSYGLRPGSDP 50

RESULT 7  
US-08-453-588-23  
Sequence 23, Application US/08453588  
Patent No. 5684145  
GENERAL INFORMATION:  
APPLICANT: Anna van der Zee, Irma Marianne van Die,  
APPLICANT: Willem Pieter Martin Hoekstra,  
APPLICANT: Josephus Theodorus Gieken.  
TITLE OF INVENTION: Carrier system against GnRH  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo No. 5684145el Patent Department  
STREET: 1300 Piccard Drive, Suite 206  
CITY: Rockville  
STATE: Maryland  
COUNTRY: U.S.A.  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,588  
FILING DATE: 30-MAY-1995

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/078,661  
FILING DATE: 16-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary E. Gormley  
REGISTRATION NUMBER: 34,409  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 258-5200  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-588-23

Query Match 59.6%; Score 62; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.025;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11  
DB 3 HWSYGLRPGS 12

RESULT 8  
US-08-521-079-23  
Sequence 23, Application US/08521079  
Patent No. 6019983  
GENERAL INFORMATION:  
APPLICANT: Anna van der Zee, Irma Marianne van Die,  
APPLICANT: Willem Pieter Martin Hoekstra,  
APPLICANT: Josephus Theodorus Gieken.  
TITLE OF INVENTION: Carrier system against GnRH  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo Pharma  
STREET: 1330 Piccard Drive  
CITY: Rockville  
STATE: Maryland  
COUNTRY: U.S.A.  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/521,079  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/078,661  
FILING DATE: 16-JUN-1993  
APPLICATION NUMBER: EPA No. 6019983 92.201.775.1  
FILING DATE: 18-JUN-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: William M. Blackstone  
REGISTRATION NUMBER: 29,772  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 258-5200  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-521-079-23



Query Match 59.6%; Score 62; DB 3; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0 025;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11  
Db 3 HWSYGLRPGS 12

RESULT 9  
US-07-690-983D-18  
Sequence 18, Application US/07690983D  
Patent No. 5403586  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J.  
APPLICANT: STEWART, Andrew G.  
APPLICANT: TSONIS, Con G.  
TITLE OF INVENTION: FUSION PROTEINS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/690,983D  
FILING DATE: 25-JUN-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU90/00373  
FILING DATE: 24-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/148 CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-690-983D-18

Query Match 59.6%; Score 62; DB 1; Length 17;  
Best Local Similarity 90.9%; Pred. No. 0 03;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 12  
Db 5 HWSYGLRPGS 15

RESULT 10  
US-07-690-983D-28  
Sequence 28, Application US/07690983D  
Patent No. 5403586  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J.  
APPLICANT: STEWART, Andrew G.  
APPLICANT: TSONIS, Con G.  
TITLE OF INVENTION: FUSION PROTEINS  
NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/690,983D  
FILING DATE: 25-JUN-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU90/00373  
FILING DATE: 24-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/148 CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-690-983D-28

Query Match 59.6%; Score 62; DB 1; Length 18;  
Best Local Similarity 90.9%; Pred. No. 0 032;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 12  
Db 6 HWSYGLRPGS 16

RESULT 11  
US-08-387-156-4  
Sequence 4, Application US/08387156  
Patent No. 5723129  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HUI P. A.  
TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/387,156  
FILING DATE: 10-FEB-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-387-156-4

Query Match 59.6%; Score 62; DB 1; Length 49;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11  
|||||  
DB 2 HWSYGLRPGS 11

RESULT 12  
US-08-694-865-4  
Sequence 4, Application US/08694865  
Patent No. 5837268  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,865  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-865-4

Query Match 59.6%; Score 62; DB 2; Length 49;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11

|||||  
DB 2 HWSYGLRPGS 11

RESULT 13  
US-08-878-748-4  
Sequence 4, Application US/08878748  
Patent No. 5969126  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P.A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/878,748  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-878-748-4

Query Match 59.6%; Score 62; DB 2; Length 49;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11  
|||||  
DB 2 HWSYGLRPGS 11

RESULT 14  
US-09-124-491-4  
Sequence 4, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-4

Query Match 59.6%; Score 62; DB 3; Length 49;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11  
DB 2 HWSYGLRPGS 11

RESULT 15  
US-08-387-156-10  
Sequence 10, Application US/08387156  
Patent No. 5723129  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/387,156  
FILING DATE: 10-FEB-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-387-156-10

Query Match 59.6%; Score 62; DB 1; Length 544;  
Best Local Similarity 100.0%; Pred. No. 0.85;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11  
DB 495 HWSYGLRPGS 504

Search completed: March 2, 2001, 10:54:36  
Job time: 601 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 2, 2001, 11:00:47 ; Search time 97.13 Seconds  
(without alignments)  
20.514 Million cell updates/sec

Title: US-09-306-689-4  
Perfect score: 104  
Sequence: 1 XHMSYGLRPGSSPPPC 17

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

SPTREMBL.15:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	55.8	91	13	09PRH0
2	53	51.0	112	12	088020
3	52	50.0	87	13	09Y126
4	52	50.0	95	13	073812
5	52	50.0	99	13	09IA10
6	52	50.0	112	12	004273
7	52	50.0	112	12	088050
8	52	50.0	112	12	088057
9	52	50.0	112	12	088058
10	52	50.0	112	12	088064
11	52	50.0	112	12	088068
12	52	50.0	112	12	088125
13	52	50.0	112	12	088131
14	52	50.0	112	12	088137
15	52	50.0	112	12	088145
16	52	50.0	112	12	089827
17	52	50.0	112	12	007393
18	52	50.0	112	12	085604
19	52	50.0	112	12	089552

20	51.5	49.5	678	4	094850	094850 homo sapien
21	51	49.0	381	5	09VMD2	09VMD2 dtrosophila
22	51	49.0	862	4	09NT23	09NT23 homo sapien
23	51	49.0	1049	4	094957	094957 homo sapien
24	50	48.1	90	13	09IAU2	09IAU2 rana dybows
25	50	48.1	112	12	076623	076623 human immun
26	50	48.1	112	12	P88144	P88144 human immun
27	50	48.1	112	12	09Q1S9	09Q1S9 human immun
28	50	48.1	112	12	09Q1S8	09Q1S8 human immun
29	50	48.1	112	12	09Q1S7	09Q1S7 human immun
30	50	48.1	112	12	09Q1S6	09Q1S6 human immun
31	50	48.1	112	12	09PX03	09PX03 human immun
32	50	48.1	1175	6	09TV66	09TV66 oryctolagus
33	50	48.1	1198	11	09JKA7	09JKA7 rattus norv
34	50	48.1	1203	4	09Y3Q4	09Y3Q4 homo sapien
35	50	48.1	1203	4	09UMQ7	09UMQ7 homo sapien
36	49.5	47.6	260	5	09V435	09V435 dtrosophila
37	49.5	47.6	653	11	P97543	P97543 rattus norv
38	49	47.1	754	5	09N5M3	09N5M3 caenorhabdi
39	49	47.1	1051	5	09Y1S0	09Y1S0 dtrosophila
40	49	47.1	1616	4	01S054	01S054 homo sapien
41	49	47.1	2129	5	09U1R8	09U1R8 caenorhabdi
42	48.5	46.6	101	12	09QW86	09QW86 human immun
43	48.5	46.6	208	5	09VW18	09VW18 dtrosophila
44	48.5	46.6	645	4	09P282	09P282 homo sapien
45	48	46.2	300	12	089582	089582 bovine hepr

## ALIGNMENTS

RESULT	ID	Q9PRH0	PRELIMINARY:	PRT:	91 AA.
AC	Q9PRH0				
DT	01-MAY-2000	(TREMBLrel. 13, Created)			
DT	01-MAY-2000	(TREMBLrel. 13, last sequence update)			
DT	01-JUN-2000	(TREMBLrel. 14, last annotation update)			
DE	PREPRO-MGNRH PRECURSOR.				
OS	Anguilla japonica (Japanese eel).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
CC	Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;				
OC	Anguillidae; Anguilla.				
OX	NCBI_TaxID=7937;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=BRAIN;				
RA	Okubo K., Suetake H., Aida K.;				
RT	"Expression of two gonadotropin-releasing hormone (GNRH) precursor				
RT	genes in various tissues of the Japanese eel and evolution of GNRH.";				
RL	Zool. Sci. 16:471-478(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Okubo K., Suetake H., Aida K.;				
RT	"A splicing variant for the prepro-mammalian gonadotropin-releasing				
RT	hormone (prepro-mGNRH) mRNA is present in the brain and various				
RT	peripheral tissues of the Japanese eel.";				
RL	Zool. Sci. 16:645-651(1999).				
DR	EMBL: AB026989; BAA82608.1; -				
DR	EMBL: AB026991; BAA83597.1; -				
DR	INTERPRO: IPR002012; -				
DR	PFAM: PF00446; GNRH; 1.				
DR	PROSITE: PS00473; GNRH; 1.				
KW	SIGNAL.				
FT	SIGNAL	1	22	POTENTIAL.	
FT	CHAIN	23	32	MGNRH.	
FT	CHAIN	33	91	GNRH ASSOCIATED PEPTIDE.	
SO	SEQUENCE	91 AA.	9893 MW.	BAI5C9DC08434A7B CRC64;	

Query Match 55.8%; Score 58; DB 13; Length 91;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||  
 DB 24 HWSYGLRPG 32

RESULT 2  
 ID 088020 PRELIMINARY; PRT; 112 AA.

DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, last annotation update)  
 DE (MM316ZA) DNA, COMPLETE CDS'S.  
 GN VPX.  
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
 OC Viruses; Retrovirda; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11723;  
 OX (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94016835; Pubmed=8411355;  
 RA Kodana T., Mori K., Kawahara T., Ringler D.J., Destroiers R.C.;  
 RT Analysis of simian immunodeficiency virus sequence variation in  
 RL tissues of rhesus macaques with simian AIDS.\*;  
 J. Virol. 67:6522-6534(1993).  
 DE EMBL: L22807; AAA47688.1; -;  
 DR INTERPRO: IPR000012; -;  
 DR PRAM: PR00522; VPR: 1;  
 DR PRINTS: PR00444; HIVVPRPX.  
 SQ SEQUENCE 112 AA; 12945 MW; A933BF982AFD7BE8 CRC64;

Query Match 51.0%; Score 53; DB 12; Length 112;  
 Best Local Similarity 60.0%; Pred. No. 1.4;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16  
 |||||  
 DB 94 HGAEGWRPSPPPPP 108

RESULT 3  
 ID 09126 PRELIMINARY; PRT; 87 AA.  
 AC 09126;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)  
 DE (FRAGMENT).  
 OS Sparus aurata (Gilthead sea bream).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;  
 OC Sparidae; Sparus.  
 NCBI\_TaxID=8175;  
 OX (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Ovary;  
 RA Nabissi M.;  
 RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
 DE EMBL: AF046801; AAD02427.1; -;  
 DR INTERPRO: IPR002012; -;  
 DR PRAM: PR00446; GNRH: 1;  
 DR PROSITE: PS00473; GNRH: 1;  
 KM Hormone: Amidation.  
 FT NON\_TER 1 87  
 FT NON\_TER 87 87  
 SQ SEQUENCE 87 AA; 9871 MW; 002463533D96782A CRC64;

Query Match 50.0%; Score 52; DB 13; Length 87;  
 Best Local Similarity 88.9%; Pred. No. 1.5;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||  
 DB 22 HWSYGLSPG 30

RESULT 4  
 ID 073812 PRELIMINARY; PRT; 95 AA.

DT 01-AUG-1998 (TREMblrel. 07, Created)  
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).  
 OS Morone saxatilis (Striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;  
 OC Moronidae; Morone.  
 NCBI\_TaxID=34816;  
 OX (1)  
 RP SEQUENCE FROM N.A.  
 RA Chow M.M., Kight K.E., Gotthilf Y., Alok D., Zohar Y.;  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC EMBL: AF056314; AAD03817.1; -;  
 DR INTERPRO: IPR002012; -;  
 DR PRAM: PR00446; GNRH: 1;  
 DR PROSITE: PS00473; GNRH: 1;  
 DR PRODOM: PD00581; -; 1;  
 KM Hormone: Amidation.  
 SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match 50.0%; Score 52; DB 13; Length 95;  
 Best Local Similarity 88.9%; Pred. No. 1.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||  
 DB 24 HWSYGLSPG 32

RESULT 5  
 ID 091A10 PRELIMINARY; PRT; 99 AA.  
 AC 091A10;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)  
 DE GONADOTROPIN RELEASING HORMONE SEABREAM ISOFORM.  
 OS Dicentrarchus labrax (European sea bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;  
 OC Moronidae; Dicentrarchus.  
 NCBI\_TaxID=13489;  
 OX (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanny S.,  
 RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;  
 RT Differential expression of three different prepro-GNRH  
 RT (gonadotropin-releasing hormone) messengers in the brain of the  
 RT European sea bass (Dicentrarchus labrax).  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN:  
RA Zmora N., Zohar Y., Elizur A.:  
RT "3 GnRH form in the seabass Dicentrarchus labrax."  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF224279; AAF62898.1; -  
SQ SEQUENCE 99 AA; 10758 MW; ECA8AEEC3CC02904 CRC64;

Query Match 50.0%; Score 52; DB 13; Length 99;  
Best Local Similarity 88.9%; Pred. No. 1.7;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
| | | | | | | |  
DB 28 HWSYGLSPG 36

RESULT 6  
ID 004273 PRELIMINARY; PRT; 112 AA.  
AC 004273;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE VPX PROTEIN (X ORF PROTEIN).  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Beveniste R.E., Heidecker G., Greenwood J., Gonda M.A.;  
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.  
RC -I- MISCELLANEOUS: THIS IS A CAPTIVE MACAQUE NEMESTRINA ISOLATE.  
DR EMBL: M32741; AAA91933.1; -  
DR INTERPRO: IPR000012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVFX.  
KW AIDS.  
SQ SEQUENCE 112 AA; 12904 MW; 7E1564F45AFD7BEA CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;  
Best Local Similarity 60.0%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16  
| : | | | | | | | |  
DB 94 HGAGGMRPGPPPP 108

RESULT 7  
ID 088050 PRELIMINARY; PRT; 112 AA.  
AC 088050;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE (MM316x) DNA, COMPLETE CDS.S.  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-94016835; PubMed-8411355;  
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
RT "Analysis of simian immunodeficiency virus sequence variation in  
RT tissues of rhesus macaques with simian AIDS."  
RL J. Virol. 67:6522-6534(1993).  
DR EMBL: L22804; AAA47679.1; -  
DR PRINTS: PR00444; HIVPRVFX.  
DR INTERPRO: IPR000012; -  
SQ SEQUENCE 112 AA; 12846 MW; 7FC007A3BAFD7BEA CRC64;

DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVFX.  
SQ SEQUENCE 112 AA; 12872 MW; 44BE094E1955B60 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;  
Best Local Similarity 60.0%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16  
| : | | | | | | | |  
DB 94 HGAGGMRPGPPPP 108

RESULT 8  
ID 088057 PRELIMINARY; PRT; 112 AA.  
AC 088057;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE (MM316x) DNA, COMPLETE CDS.S.  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-94016835; PubMed-8411355;  
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
RT "Analysis of simian immunodeficiency virus sequence variation in  
RT tissues of rhesus macaques with simian AIDS."  
RL J. Virol. 67:6522-6534(1993).  
DR EMBL: L22791; AAA72399.1; -  
DR INTERPRO: IPR000012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVFX.  
SQ SEQUENCE 112 AA; 12842 MW; BDB78F982AFD7BF2 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;  
Best Local Similarity 60.0%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16  
| : | | | | | | | |  
DB 94 HGAGGMRPGPPPP 108

RESULT 9  
ID 088058 PRELIMINARY; PRT; 112 AA.  
AC 088058;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE (MM316L) DNA, COMPLETE CDS.S.  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-94016835; PubMed-8411355;  
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
RT "Analysis of simian immunodeficiency virus sequence variation in  
RT tissues of rhesus macaques with simian AIDS."  
RL J. Virol. 67:6522-6534(1993).  
DR EMBL: L22792; AAA47643.1; -  
DR INTERPRO: IPR000012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVFX.  
SQ SEQUENCE 112 AA; 12846 MW; 7FC007A3BAFD7BEA CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;  
Best Local Similarity 60.0%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16  
1 : 1 1 1 1 1 1 1  
DB 94 HGAGGMRPGPPPPPP 108

RESULT 10  
ID 088064 PRELIMINARY; PRT; 112 AA.  
AC 088064;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE (MM3160) DNA, COMPLETE CDS'S.  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94016835; PubMed=8411355;  
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
RT "Analysis of simian immunodeficiency virus sequence variation in  
tissues of rhesus macaques with simian AIDS."  
RL J. Virol. 67:6522-6534(1993).  
DR EMBL: L22795; AAA47652.1; -  
DR INTERPRO: IPR000012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVPX.  
SQ SEQUENCE 112 AA; 12935 MW; D32070A4DAC1CF5 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;  
Best Local Similarity 60.0%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16  
1 : 1 1 1 1 1 1 1  
DB 94 HGAGGMRPGPPPPPP 108

RESULT 11  
ID 088068 PRELIMINARY; PRT; 112 AA.  
AC 088068;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE (MM3160) DNA, COMPLETE CDS'S.  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94016835; PubMed=8411355;  
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;  
RT "Analysis of simian immunodeficiency virus sequence variation in  
tissues of rhesus macaques with simian AIDS."  
RL J. Virol. 67:6522-6534(1993).  
DR EMBL: L22797; AAA47658.1; -  
DR INTERPRO: IPR000012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVPX.  
SQ SEQUENCE 112 AA; 12841 MW; 8B82EF96CA137BF0 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;

Best Local Similarity 60.0%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16  
1 : 1 1 1 1 1 1 1  
DB 94 HGAGGMRPGPPPPPP 108

RESULT 12  
ID 088125 PRELIMINARY; PRT; 112 AA.  
AC 088125;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VPX PROTEIN.  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90272009; PubMed=1971917;  
RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;  
RT "Sequence analysis and acute pathogenicity of molecularly cloned  
SIVSM-PB14."  
RL Nature 345:636-640(1990).  
DR EMBL: L03296; AAA47763.1; -  
DR INTERPRO: IPR000012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVPX.  
SQ SEQUENCE 112 AA; 12877 MW; BEADC7FF6CFD2EAC CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;  
Best Local Similarity 60.0%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16  
1 : 1 1 1 1 1 1 1  
DB 94 HGAGGMRPGPPPPPP 108

RESULT 13  
ID 088131 PRELIMINARY; PRT; 112 AA.  
AC 088131;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE VPX PROTEIN.  
GN VPX.  
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90272009; PubMed=1971917;  
RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;  
RT "Sequence analysis and acute pathogenicity of molecularly cloned  
SIVSM-PB14."  
RL Nature 345:636-640(1990).  
DR EMBL: L03297; AAA47771.1; -  
DR INTERPRO: IPR000012; -  
DR PFAM: PF00522; VPR; 1.  
DR PRINTS: PR00444; HIVPRVPX.  
SQ SEQUENCE 112 AA; 12851 MW; 9C3EC7FF6CFD3D14 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;  
Best Local Similarity 60.0%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;



OY 2 HWSYGLRPGSSPPP 16  
 | : | | | | | | |  
 DB 94 HGAGGWRPGPPPPP 108

## RESULT 14

O88137 PRELIMINARY: PRT: 112 AA.  
 AC O88137:  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VPX PROTEIN.  
 GN VPX.  
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90272009; PubMed=1971917;  
 RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;  
 RT "Sequence analysis and acute pathogenicity of molecularly cloned  
 RT SIVSM-PBJ14."  
 RL Nature 345:636-640(1990).  
 DR EMBL: L03295; AAB59772.1; -  
 DR INTERPRO: IPR000012; -  
 DR PFAM: PF00522; VPR: 1.  
 DR PRINTS: PR00444; HIVPRVFX.  
 SQ SEQUENCE 112 AA: 12881 MW; C36D1AE2A0FD270E CRC64;

## Query Match

Best Local Similarity 50.0%; Score 52; DB 12; Length 112;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPP 16  
 | : | | | | | | |  
 DB 94 HGAGGWRPGPPPPP 108

## RESULT 15

O88145 PRELIMINARY: PRT: 112 AA.  
 ID O88145:  
 AC O88145:  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE VPX PROTEIN.  
 GN VPX.  
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92124755; PubMed=1733112;  
 RA Novembre F.J., Hirsch V.M., McClure H.M., Fultz P.N., Johnson P.R.;  
 RT "SIV from stump-tailed macaques: molecular characterization of a  
 RT highly transmissible primate lentivirus."  
 RL Virology 186:783-787(1992).  
 DR EMBL: M83293; AAA91942.1; -  
 DR INTERPRO: IPR000012; -  
 DR PFAM: PF00522; VPR: 1.  
 DR PRINTS: PR00444; HIVPRVFX.  
 SQ SEQUENCE 112 AA: 12934 MW; ADD7F9DC45868440 CRC64;

## Query Match

Best Local Similarity 50.0%; Score 52; DB 12; Length 112;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPP 16  
 | : | | | | | | |  
 DB 94 HGAGGWRPGPPPPP 108

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:01:19 ; Search time 32.03 Seconds  
(without alignments)  
17.140 Million cell updates/sec

Title: US-09-306-689-4  
Sequence: 1 XHMSVGLRGGSPPPRC 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SWISSProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	55.8	61	1 GONI_SHEEP	Q28588 ovis aries
2	58	55.8	63	1 GONI_MESAU	O09163 mesocricetu
3	58	55.8	67	1 GONI_MACMU	P55247 macaca mula
4	58	55.8	89	1 GONI_XENIA	P45656 xenopus lae
5	58	55.8	90	1 GONI_MOUSE	P13562 mus musculu
6	58	55.8	91	1 GONI_PIG	P49921 sus scrofa
7	58	55.8	92	1 GONI_HUMAN	P01148 homo sapien
8	58	55.8	92	1 GONI_RAT	P07490 rattus norv
9	58	55.8	92	1 GONI_TUPGB	O55335 tupia glis
10	54	51.9	10	1 GONI_ALAMI	P37041 alligator m
11	54	51.9	92	1 GONI_CHICK	P37042 gallus galli
12	52	50.0	94	1 GONI_HAPBU	P51918 haplochromi
13	52	50.0	95	1 GONI_PAGMA	P70074 pagrus majo
14	52	50.0	95	1 GONI_SPAU	P51919 sparus auro
15	52	50.0	112	1 VPX_SIVM1	P05917 simian immu
16	52	50.0	112	1 VPX_SIVS	P19508 simian immu
17	51	49.0	112	1 VPX_SIVM	P05916 simian immu
18	51	49.0	112	1 VPX_SIVM	P11266 simian immu
19	50	48.1	111	1 VPX_HV2KR	Q74122 human immu
20	50	48.1	112	1 VPX_HV2CA	P24110 human immu
21	50	48.1	112	1 VPX_HV2D1	P13760 human immu
22	50	48.1	112	1 VPX_HV2RO	P06933 human immu
23	50	48.1	324	1 TTP_BOVIN	P53781 bos taurus
24	49.5	47.6	719	1 DEND_RAT	P50617 rattus norv
25	49	47.1	92	1 GONI_CAVPO	O54713 cavia porce
26	48	46.2	80	1 GONI_CLAGA	P33439 claris gar
27	48	46.2	432	1 ENV2_MOUSE	P11370 mus musculu
28	48	46.2	676	1 ENV_MLVRP	P26803 friend muri
29	48	46.2	1091	1 DIA_DROME	P48608 drosophila
30	47	45.2	315	1 ALDR_PIG	P80276 sus scrofa
31	46	44.2	309	1 HXMA_CHICK	P17277 gallus galli
32	46	44.2	395	1 TRBL_AGR16	P54913 agrobacteri
33	45	43.3	10	1 GON3_ONCKE	P20367 oncorhynch

## ALIGNMENTS

RESULT	1	STANDARD:	PRT:	61 AA.
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10
11	11	11	11	11
12	12	12	12	12
13	13	13	13	13
14	14	14	14	14
15	15	15	15	15
16	16	16	16	16
17	17	17	17	17
18	18	18	18	18
19	19	19	19	19
20	20	20	20	20
21	21	21	21	21
22	22	22	22	22
23	23	23	23	23
24	24	24	24	24
25	25	25	25	25
26	26	26	26	26
27	27	27	27	27
28	28	28	28	28
29	29	29	29	29
30	30	30	30	30
31	31	31	31	31
32	32	32	32	32
33	33	33	33	33
34	34	34	34	34
35	35	35	35	35
36	36	36	36	36
37	37	37	37	37
38	38	38	38	38
39	39	39	39	39
40	40	40	40	40
41	41	41	41	41
42	42	42	42	42
43	43	43	43	43
44	44	44	44	44
45	45	45	45	45
46	46	46	46	46
47	47	47	47	47
48	48	48	48	48
49	49	49	49	49
50	50	50	50	50
51	51	51	51	51
52	52	52	52	52
53	53	53	53	53
54	54	54	54	54
55	55	55	55	55
56	56	56	56	56
57	57	57	57	57
58	58	58	58	58
59	59	59	59	59
60	60	60	60	60
61	61	61	61	61
62	62	62	62	62
63	63	63	63	63
64	64	64	64	64
65	65	65	65	65
66	66	66	66	66
67	67	67	67	67
68	68	68	68	68
69	69	69	69	69
70	70	70	70	70
71	71	71	71	71
72	72	72	72	72
73	73	73	73	73
74	74	74	74	74
75	75	75	75	75
76	76	76	76	76
77	77	77	77	77
78	78	78	78	78
79	79	79	79	79
80	80	80	80	80
81	81	81	81	81
82	82	82	82	82
83	83	83	83	83
84	84	84	84	84
85	85	85	85	85
86	86	86	86	86
87	87	87	87	87
88	88	88	88	88
89	89	89	89	89
90	90	90	90	90
91	91	91	91	91
92	92	92	92	92
93	93	93	93	93
94	94	94	94	94
95	95	95	95	95
96	96	96	96	96
97	97	97	97	97
98	98	98	98	98
99	99	99	99	99
100	100	100	100	100
101	101	101	101	101
102	102	102	102	102
103	103	103	103	103
104	104	104	104	104
105	105	105	105	105
106	106	106	106	106
107	107	107	107	107
108	108	108	108	108
109	109	109	109	109
110	110	110	110	110
111	111	111	111	111
112	112	112	112	112
113	113	113	113	113
114	114	114	114	114
115	115	115	115	115
116	116	116	116	116
117	117	117	117	117
118	118	118	118	118
119	119	119	119	119
120	120	120	120	120
121	121	121	121	121
122	122	122	122	122
123	123	123	123	123
124	124	124	124	124
125	125	125	125	125
126	126	126	126	126
127	127	127	127	127
128	128	128	128	128
129	129	129	129	129
130	130	130	130	130
131	131	131	131	131
132	132	132	132	132
133	133	133	133	133
134	134	134	134	134
135	135	135	135	135
136	136	136	136	136
137	137	137	137	137
138	138	138	138	138
139	139	139	139	139
140	140	140	140	140
141	141	141	141	141
142	142	142	142	142
143	143	143	143	143
144	144	144	144	144
145	145	145	145	145
146	146	146	146	146
147	147	147	147	147
148	148	148	148	148
149	149	149	149	149
150	150	150	150	150
151	151	151	151	151
152	152	152	152	152
153	153	153	153	153
154	154	154	154	154
155	155	155	155	155
156	156	156	156	156
157	157	157	157	157
158	158	158	158	158
159	159	159	159	159
160	160	160	160	160
161	161	161	161	161
162	162	162	162	162
163	163	163	163	163
164	164	164	164	164
165	165	165	165	165
166	166	166	166	166
167	167	167	167	167
168	168	168	168	168
169	169	169	169	169
170	170	170	170	170
171	171	171	171	171
172	172	172	172	172
173	173	173	173	173
174	174	174	174	174
175	175	175	175	175
176	176	176	176	176
177	177	177	177	177
178	178	178	178	178
179	179	179	179	179
180	180	180	180	180
181	181	181	181	181
182	182	182	182	182
183	183	183	183	183
184	184	184	184	184
185	185	185	185	185
186	186	186	186	186
187	187	187	187	187
188	188	188	188	188
189	189	189	189	189
190	190	190	190	190
191	191	191	191	191
192	192	192	192	192
193	193	193	193	193
194	194	194	194	194
195	195	195	195	195
196	196	196	196	196
197	197	197	197	197
198	198	198	198	198
199	199	199	199	199
200	200	200	200	200
201	201	201	201	201
202	202	202	202	202
203	203	203	203	203
204	204	204	204	204
205	205	205	205	205
206	206	206	206	206
207	207	207	207	207
208	208	208	208	208
209	209	209	209	209
210	210	210	210	210
211	211	211	211	211
212	212	212	212	212
213	213	213	213	213
214	214	214	214	214
215	215	215	215	215

FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).  
 FT NON\_TER 61  
 SO SEQUENCE 61 AA: 6828 MW: 63962A1AE319B8FO CRC64:

Query Match 55.8%; Score 58; DB 1; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 Db 2 HWSYGLRPG 10

## RESULT 2

CON1\_MESAU STANDARD; PRT; 63 AA.  
 ID CON1\_MESAU  
 AC 009163;  
 DT 15-DEC-1998 (rel. 37, Created)  
 DT 15-DEC-1998 (rel. 37, Last sequence update)  
 DT 30-MAY-2000 (rel. 39, Last annotation update)  
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
 (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
 HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]  
 DE (FRAGMENT).  
 GN GNRH1 OR GNRH OR LHRH.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Jansen H.T., Stevens P.J., Zeidler P., Lehman M.N.;  
 Submitted (MAR-1987) to the EMBL/GenBank/DDI databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 HORMONES.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

CC EMBL: U91938; AAB51302.1; -  
 DR INTERPRO: IPR002012; -  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta.  
 FT NON\_TER 1 1  
 FT CHAIN 1 >63  
 FT PEPTIDE 1 10  
 FT PEPTIDE 14 >63  
 FT ACT\_SITE 3 3  
 FT MOD\_RES 1 1  
 FT MOD\_RES 10 10  
 FT NON\_TER 63 63  
 FT SEQUENCE 63 AA: 7370 MW: FC94995676F77180 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 Db 2 HWSYGLRPG 10

## RESULT 3

CON1\_MACMU STANDARD; PRT; 67 AA.  
 ID CON1\_MACMU  
 AC P55247;  
 DT 01-OCT-1996 (rel. 34, Created)  
 DT 01-OCT-1996 (rel. 34, Last sequence update)  
 DT 30-MAY-2000 (rel. 39, Last annotation update)  
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
 (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
 HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]  
 DE (FRAGMENT).  
 GN GNRH1 OR GNRH OR LHRH.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-HYPOTHALAMUS;  
 RX MEDLINE-95124501; PubMed-7545971;  
 RA Ma Y.J., Costa M.E., Ojeda S.R.;  
 RT "Developmental expression of the genes encoding transforming growth  
 factor alpha and its receptor in the hypothalamus of female rhesus  
 macaques."  
 RT Neuroendocrinology 60:346-359(1994).  
 RL Neuroendocrinology 60:346-359(1994).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 HORMONES.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

CC EMBL: S75918; AAB33096.1; -  
 DR INTERPRO: IPR002012; -  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal.  
 FT NON\_TER 1 1  
 FT SIGNAL 1 5  
 FT CHAIN 6 >67  
 FT PEPTIDE 6 15  
 FT PEPTIDE 19 >67  
 FT ACT\_SITE 8 8  
 FT MOD\_RES 6 6  
 FT MOD\_RES 15 15  
 FT NON\_TER 67 67  
 FT SEQUENCE 67 AA: 7573 MW: 505394DAA261A3F2 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 0.17;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 Db 7 HWSYGLRPG 15

```

RESULT 4
GON1_XENLA STANDARD: PRT: 89 AA.
AC P45656.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FOREBRAIN;
RA MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Wray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28040; AAA49728.1;
DR INTERPRO; IPR002012;
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 89 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT CHAIN 37 89 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
FT PEPTIDE 37 85 PEPTIDE.
FT MOD_RES 24 24 GNRH-ASSOCIATED PEPTIDE I (GAP).
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SO SEQUENCE 89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33

RESULT 5
GON1_MOUSE STANDARD: PRT: 90 AA.
AC P13562;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
DE I].
OS GNRH1 OR GNRH.
GN Mus musculus (Mouse).

```

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87069928; PubMed=3024317;
RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA Phillips H.S., Nikolics K., Seeburg P.H.;
RT "A deletion truncating the gonadotropin-releasing hormone gene is
RT responsible for hypogonadism in the hpg mouse."
RL Science 234:1366-1371(1986).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14872; AAA37717.1;
DR MED; MG1:95789; GNRH.
DR INTERPRO; IPR002012;
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placental; Signal.
FT SIGNAL 1 21
FT CHAIN 22 90 PROGONADOLIBERIN I.
FT PEPTIDE 22 31 GONADOLIBERIN I.
FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
SO SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 23 HWSYGLRPG 31

RESULT 6
GON1_PIG STANDARD: PRT: 91 AA.
AC P49921;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HYPOPHALAMUS;
RA Weesner G.D., Mattern R.L., Becker B.A.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 24-33.
RA MEDLINE=72114303; PubMed=4946067;
RA Baba Y., Matsuo H., Schally A.V.;

```

RT	FT	Structure of the porcine LH- and FSH-releasing hormone. II.
RT	RT	Confirmation of the proposed structure by conventional sequential
RT	RT	analyses. ",
RL	RL	Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN	RN	[3]
RP	RP	SYNTHESIS OF GONADOLIBERIN.
RX	RX	MEDLINE=72053376; PubMed=4942726;
RA	RA	Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT	RT	"Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RL	RL	phase method. ",
RL	RL	Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN	RN	[4]
RP	RP	SYNTHESIS OF GONADOLIBERIN.
RX	RX	MEDLINE=72117544; PubMed=4946275;
RA	RA	Baba Y., Arimura A., Schally A.V.;
RT	RT	"On the tyrosine residue in porcine LH and FSH-releasing hormone. ",
RL	RL	Biochem. Biophys. Res. Commun. 45:483-487(1971).
CC	CC	-I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC	CC	THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC	CC	HORMONES.
CC	CC	-I- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC	CC	-----
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	CC	the European Bioinformatics Institute. There are no restrictions on its
CC	CC	use by non-profit institutions as long as its content is in no way
CC	CC	modified and this statement is not removed. Usage by and for commercial
CC	CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	CC	-----
DR	DR	EMBL, L32864; AAA31066.1; .
DR	DR	PIR, A01411; RHPEG.
DR	DR	INTERPRO: IPR002012; .
DR	DR	PFAM: PF00446; GnRH; 1.
DR	DR	PROSITE: PS00473; GnRH; 1.
KW	KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
FT	FT	Placenta; Signal.
FT	FT	SIGNAL 1 23
FT	FT	CHAIN 24 91
FT	FT	PEPTIDE 24 33
FT	FT	PEPTIDE 34 91
FT	FT	ACT_SITE 26 26
FT	FT	MOD_RES 24 24
FT	FT	MOD_RES 33 33
FT	FT	SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
QY	QY	Query Match 55.8%; Score 58; DB 1; Length 91;
DB	DB	Best Local Similarity 100.0%; Pred. No. 0.23;
		Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
		2 HWSYGLRPG 10
		25 HWSYGLRPG 33
RESULT 7		
GONI_HUMAN	STANDARD:	PRT; 92 AA.
ID GONI_HUMAN	AC P01148;	
DT 21-JUL-1986	(Rel. 01, Created)	
DT 01-APR-1988	(Rel. 07, Last sequence update)	
DT 30-MAY-2000	(Rel. 39, Last annotation update)	
DE PROGONADOLIBERIN I PRECURSOR (CONTAINS GONADOLIBERIN I (LHRH I)		
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING		
DE HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN) (GNRH-ASSOCIATED		
DE PEPTIDE I).		
GN GNRI OR GNRH OR LHRH.		
OS Homo sapiens (Human).		
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;		
OC Mammalia, Euteria, Primates, Catarrhini, Homnidae, Homo.		
NC [1]		

```

RX SEQUENCE FROM N.A. Pubmed-2671939:
RA Hayflick J.S., Adelman J.P., Seeburg P.H.:
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RN hormone gene.":
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [12]
RP SEQUENCE FROM N.A. Pubmed-2867548:
RX MEDLINE=66094338;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.:
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RN of gonadotropin-releasing hormone and prolactin release-inhibiting
RL factor in human and rat.":
RP Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [13]
RP SEQUENCE FROM N.A. Pubmed-609051:
RX MEDLINE=85012739;
RA Seeburg P.H., Adelman J.P.:
RT "Characterization of cDNA for precursor of human luteinizing hormone
RN releasing hormone.":
RL Nature 311:666-668(1984).
RN [14]
RP SEQUENCE OF 24-33.
RX MEDLINE=83126573; Pubmed=6760865;
RA Tan L., Rousseau P.:
RT "The chemical identity of the immunoreactive LHRH-like peptide
RN biosynthesized in the human placenta.":
CC Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -I- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS),
CC LUTREPUSE OR LUTRELEF (FERRING PHARMACEUTICALS) AND RELISORM
CC (SERONO).
CC -I- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; X01059; CAA25526.1; -
DR EMBL; M12578; AAA35916.1; -
DR EMBL; X15215; CAA33285.1; -
DR PIR; A01410; RHUG.
DR PIR; A26173; A26173.
DR PIR; S05308; S05308.
DR MIM; 152760; -
DR INTERPRO; IPR002012; -
DR PFM; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT CONFLICT 16 16
FT CONFLICT 16 16
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

```

Db 25 HWSYGLRPG 33

```

RESULT 8
GONL_RAT STANDARD: PRT: 92 AA.
ID GONL_RAT
AC P07490:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)
  (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
  HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
  I).
GN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-86094338; PubMed-2867548;
  Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
  "Isolation of the gene and hypothalamic cDNA for the common precursor
  of gonadotropin-releasing hormone and prolactin release-inhibiting
  factor in human and rat.";
  Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE-893844661; PubMed-2476669;
  Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
  "The rat gonadotropin-releasing hormone: SH locus: structure and
  hypothalamic expression.";
  Mol. Endocrinol. 3:1257-1262(1989).
RL [3]
RN SEQUENCE FROM N.A.
RP TISSUE-THYMUS.
RC MEDLINE-93105480; PubMed-1468115;
  Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
  "Thymocytes express a mRNA that is identical to hypothalamic
  luteinizing hormone-releasing hormone mRNA.";
  Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE-HEART.
RA MEDLINE-87149087; PubMed-3547652;
  Adelman J.P., Bond C.T., Douglass J., Herbert E.;
  "Two mammalian genes transcribed from opposite strands of the same
  DNA locus.";
  Science 235:1514-1517(1987).
RL -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
  HORMONES.
CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -
DR EMBL; M12579; AAA41263.1; -
DR EMBL; M31670; AAA41264.1; -
DR EMBL; M15527; AAA42141.1; ALT-SEQ.
DR EMBL; M15529; AAA42139.1; -
DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR PIR; B26173; RHRTG.
DR PIR; A48410; A48410.
DR INTERPRO; IPR002012; -

```

```

DR PFAM; PF00446; GNRH. 1.
DR PROSITE; PS00473; GNRH. 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
  Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33

RESULT 9
GONL_TUPGB STANDARD: PRT: 92 AA.
ID GONL_TUPGB
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)
  (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
  HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRH1 OR GNRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Scandentia; Tupalidae; Tupala.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HYPOTHALAMUS;
  MEDLINE-97079639; PubMed-8921350;
  Kaestlen T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
  Fernald R.D.;
  "Characterization of two new preproGNRH mRNAs in the tree shrew:
  first direct evidence for mesencephalic GNRH gene expression in a
  placental mammal.";
  Gen. Comp. Endocrinol. 104:7-19(1996).
RL -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
  HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U63326; AAB16837.1; -
DR INTERPRO; IPR002012; -
DR PFAM; PF00446; GNRH. 1.
DR PROSITE; PS00473; GNRH. 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
  Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
BY SIMILARITY.
PROGNADOLIBERIN I.
GONADOLIBERIN I.
GNRH-ASSOCIATED PEPTIDE I.
APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.

```

FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY  
FT MOD\_RES 33 33 SIMILARITY).  
FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY  
FT MOD\_RES 33 33 SIMILARITY).  
SO SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;

Query Match 55.8%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
Db 25 HWSYGLRPG 33

RESULT 10  
GONI\_ALAMI STANDARD; PRT; 10 AA.  
AC P37041; P20407;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE GONADOLIBERIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I)  
DE (LULIBERIN I).  
OS Alligator mississippiensis (American alligator).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.  
RN [1]  
RP SEQUENCE.  
RC TISSUE-BRAIN;  
RX MEDLINE=91352338; PubMed=1882082;  
RA Lavey D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,  
Lance V., Stanson P., Rivier J.E., Sherwood N.M.;  
RT Primary structure of two forms of gonadotropin-releasing hormone  
from brains of the American alligator (Alligator mississippiensis).  
RL Regul. Pept. 33:105-116(1991).  
CC - FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
CC - SIMILARITY: BELONGS TO THE GNRH FAMILY.  
DR INTERPRO: IPR002012;  
DR PIR: A60066; RHA01.  
DR PRAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KW Hormone; Amidation; Hypothalamus.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 10 10 AMIDATION.  
SO SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 51.9%; Score 54; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.085;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 11  
GONI\_CHICK STANDARD; PRT; 92 AA.  
AC P37042; P20407;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)  
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
DE (HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-WHITE LEHORN;  
RX MEDLINE=94059355; PubMed=7902095;  
RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;  
RT "Characterization of the chicken preprogonadotropin-releasing  
RT hormone-I gene.";  
RL J. Mol. Endocrinol. 11:19-29(1993).  
RN [2]  
RP SEQUENCE OF 24-33.  
RC TISSUE-HYPOTHALAMUS;  
RX MEDLINE=82265778; PubMed=7050119;  
RA King J.A., Millar R.P.;  
RT "Structure of chicken hypothalamic luteinizing hormone-releasing  
RT hormone. II. Isolation and characterization.";  
RL J. Biol. Chem. 257:10729-10732(1982).  
RN [3]  
RP SEQUENCE OF 24-33.  
RC TISSUE-HYPOTHALAMUS;  
RA King J.A., Millar R.P.;  
RT "Structure of avian hypothalamic gonadotropin-releasing hormone.";  
RL S. Afr. J. Sci. 78:124-125(1982).  
RN [4]  
RP SYNTHESIS OF 24-33.  
RX MEDLINE=82265777; PubMed=7050118;  
RA King J.A., Millar R.P.;  
RT "Structure of chicken hypothalamic luteinizing hormone-releasing  
RT hormone. I. Structural determination on partially purified  
RT material.";  
RL J. Biol. Chem. 257:10722-10728(1982).  
CC - FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
CC - SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC - THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X69491; CAA49246.1;  
DR PTR: S33507; S33507.  
DR INTERPRO: IPR002012;  
DR PRAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
FT SIGNAL. 1 23  
FT CHAIN 24 92 PROGONADOLIBERIN I.  
FT PEPTIDE 24 92 GONADOLIBERIN I.  
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.  
FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
SO SEQUENCE 92 AA; 10206 MW; 61ABE7EBAF508B6A CRC64;

Query Match 51.9%; Score 54; DB 1; Length 92;  
Best Local Similarity 88.9%; Pred. No. 0.75;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
Db 25 HWSYGLRPG 33

RESULT 12  
GONI\_HAPBU STANDARD; PRT; 94 AA.  
AC P51918; 093387;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)



DE (LH-RH I) (LULIBERIN I).  
 GN \*GNRH1.  
 OS Haplochromis butoni.  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;  
 CC Clichidae; Astatotilapia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-95396797; PubMed-7667296;  
 RA White S.A., Kaasen T.L., Bond C.T., Adelman J.P., Fernald R.D.;  
 RT "Three gonadotropin-releasing hormone genes in one organism suggest  
 RT novel roles for an ancient peptide."  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99061842; PubMed-9843638;  
 RA White R.B., Fernald R.D.;  
 RT "Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression  
 RT reveals a distinct origin for GNRH-containing neurons in the  
 RT midbrain."  
 RL Gen. Comp. Endocrinol. 112:322-329(1998).  
 RN [3]  
 RP SEQUENCE OF 23-32.  
 RC TISSUE-PITUITARY;  
 RX MEDLINE-95372591; PubMed-7644702;  
 RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,  
 RA White S.A., Francis R.C., Fernald R.D., Licht P., Wardy C.,  
 RA Sherwood N.M.;  
 RT "Primary structure of solitary form of gonadotropin-releasing hormone  
 RT (GNRH) in cichlid pituitary; three forms of GNRH in brain of cichlid  
 RT and pumpkinseed fish."  
 RL Regul. Pept. 57:43-53(1995).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE  
 CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-  
 CC GONADAL AXIS.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS  
 CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOPHYSAL AXONS.  
 CC -1- MASS SPECTROMETRY: MW-1113.9; METHOD-MALDI; RANGE-23-32.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U31865; AAC59691.1; -  
 DR EMBL: AF076961; AAC27716.1; -  
 DR INTERPRO: IPR002012; -  
 DR PFM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Multigene family.  
 FT SIGNAL 1 22  
 FT CHAIN 1 22 PROGONADOLIBERIN I.  
 FT PEPTIDE 23 94 GONADOLIBERIN I.  
 FT MOD\_RES 36 94 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).  
 FT MOD\_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP).  
 FT CONFLICT 86 94 ENGRITFK --> KMDTGHSHNERFL (IN REF. 1).  
 SO SEQUENCE 94 AA; 10362 MW; E57DBA8333278D7 CRC64;

Query Match 50.0%; Score 52; DB 1; Length 94;  
 Best Local Similarity 88.9%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||  
 DB 24 HWSYGLSPG 32

RESULT 13  
 GONL\_PGMA STANDARD; PRT; 95 AA.  
 ID GONL\_PGMA  
 AC P70074;  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)  
 DE (LH-RH I) (LULIBERIN I).  
 OS Pagrus major (Red sea bream) (Chrysophrys major).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 CC Sparidae; Chrysophrys.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Okuzawa K., Graneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D86582; BAA13129.1; -  
 DR INTERPRO: IPR002012; -  
 DR PFM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Multigene family.  
 FT SIGNAL 1 23  
 FT CHAIN 1 23 POTENTIAL.  
 FT PEPTIDE 24 95 GONADOLIBERIN I.  
 FT MOD\_RES 37 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY  
 FT SIMILARITY).  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP)  
 FT (BY SIMILARITY).  
 SO SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;

Query Match 50.0%; Score 52; DB 1; Length 95;  
 Best Local Similarity 88.9%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||  
 DB 25 HWSYGLSPG 33

RESULT 14  
 GONL\_SPAU STANDARD; PRT; 95 AA.  
 ID GONL\_SPAU  
 AC P51919;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)  
 DE (LH-RH I) (LULIBERIN I) (SBGNRH).  
 OS Sparus aurata (Gilthead sea bream).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 CC Sparidae; Sparus.

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=95268449; PubMed-7749463;
RA Gotlib Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-
RL releasing hormone from the gilthead seabream (Sparus aurata).";
RM Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE=BRAIN;
RX MEDLINE=95083645; PubMed-7991588;
RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
RL brains of one species.";
RM Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
RN [3]
RP FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
RC -1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U30320; AAA75469.1; -
DR INTERPRO: IPR002012; -
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KV Cleavage on pair of basic residues: Hormone; Amidation: Hypothalamus;
KT Signal, Multigene family.
FT SIGNAL 1 25
FT CHAIN 1 25
FT PEPTIDE 26 35
FT PEPTIDE 26 35
FT MOD_RES 26 35
FT MOD_RES 35 35
SQ SEQUENCE 95 AA; 10753 MW; 49313FDE6DB87DA CRC64;

Query Match
Best Local Similarity 50.0%; Score 52; DB 1; Length 95;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 27 HWSYGLSPG 35

RESULT 15
VPX_STVM1 STANDARD; PRT; 112 AA.
AC P05917;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE VPX PROTEIN (X ORF PROTEIN).
GN VPX.
OS Simian immunodeficiency virus (Mm142-83 isolate) (STV-MAC).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87287230; PubMed-3649576;
RA Chakrabarti L., Guyader M., Alizon M., Daniel M.D., Desrosiers R.C.,
RA Hjalals P., Sonigo P.;
RT "Sequence of simian immunodeficiency virus from macaque and its
RT relationship to other human and simian retroviruses.";
RL Nature 328:543-547(1987).
CC -1- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.

```

```

CC [1]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Y00277; CAA68382.1; -
DR PIR: D28887; ASLTX3.
DR HIV: M16403; VPXMM142.
DR INTERPRO: IPR000012; -
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVPX.
KW AIDS.
SQ SEQUENCE 112 AA; 12906 MW; 46F564F45AFD960 CRC64;

Query Match
Best Local Similarity 50.0%; Score 52; DB 1; Length 112;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPP 16
Db 94 HGAGGWRGPPPPPP 108

```

Search completed: March 2, 2001, 11:01:20  
 Job time: 435 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:55:39 ; Search time 56.87 Seconds  
(without alignments)  
20.297 Million cell updates/sec

Title: US-09-306-689-4

Perfect score: 104

Sequence: 1 XHMSYGLRPGSSPPPC 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	55.8	10	1	RHPCG gonadoliberin - pi
2	58	55.8	10	1	RHSHG gonadoliberin - sh
3	58	55.8	67	2	I78541 gonadoliberin prec
4	58	55.8	89	2	I51423 gonadoliberin prec
5	58	55.8	90	1	RHMSG gonadoliberin prec
6	58	55.8	92	1	RHHUG gonadoliberin prec
7	58	55.8	92	1	RHRTG gonadoliberin prec
8	58	55.8	92	1	RHAQ1 gonadoliberin I -
9	54	51.9	92	2	I50644 gonadoliberin I pr
10	52	50.0	98	2	I50739 gonadotropin-relea
11	52	50.0	106	2	A30996 orf-X protein - si
12	52	50.0	112	1	ASLJX3 vpu protein - simi
13	52	50.0	112	2	S03066 gene X protein - h
14	51	49.0	112	1	ASLJST vpu protein - simi
15	51	49.0	862	2	T46289 hypothetical prote
16	50	48.1	112	1	ASLJX2 vpu protein - huma
17	50	48.1	112	1	ASLJX2 vpu protein - huma
18	50	48.1	112	2	SI2155 vpx protein - huma
19	50	48.1	112	2	S53094 vpx protein - huma
20	49	47.1	112	2	E83400 pyrrolidinoline q
21	49	47.1	2129	1	RHJDS hypothetical prote
22	48	46.2	80	1	RHJDS gonadoliberin I pr
23	48	46.2	300	2	S41171 transrepressor pro
24	48	46.2	432	2	A25483 env polypeptide, r
25	48	46.2	676	1	VCWEPV env polypeptide, r
26	48	46.2	833	2	G75621 arylesterase/monox
27	48	46.2	969	2	T15446 hypothetical prote
28	48	46.2	1091	2	T13170 diaphanous protein
29	47	45.2	199	2	SI4981 exonin class I (

30	47	45.2	316	2	A59021 aldehyde reductase
31	47	45.2	320	2	T48188 aldose reductase-1
32	47	45.2	367	2	T01751 gibberellin 20-oxi
33	47	45.2	1438	2	B71610 WD40 WEB-1 homolog
34	46	44.2	30	2	S59482 hydroxyproline-ric
35	46	44.2	157	2	T02034 early light-induce
36	46	44.2	181	2	T35751 hypothetical prote
37	46	44.2	256	2	S74928 hypothetical prote
38	46	44.2	309	1	S09257 homeotic protein H
39	46	44.2	422	2	T09742 drought-induced pr
40	46	44.2	546	2	T02029 DNA-binding protei
41	46	44.2	1373	2	JE0095 gastric mucin MUC5
42	45	43.3	10	2	A21114 gonadoliberin - ch
43	45	43.3	74	2	T51092 gonadotropin relea
44	45	43.3	82	2	T51180 gonadotropin-relea
45	45	43.3	82	2	SI1365 gonadotropin-relea

#### ALIGNMENTS

```

RESULT 1
RHPGC gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A>Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A>Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 55.8%; Score 58; DB 1; Length 10;
Best local similarity 100.0%; Pred. No. 0.054;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSYGLRPG 10
DB 2 HMSYGLRPG 10

RESULT 2
RHSNG gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780

```

A:Molecule type: protein  
A:Residues: 1-10 <BUR>  
A:Note: the natural and synthetic hormones have the same biological activity  
C:Comment: this hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 55.8%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 2 HWSYGLRPG 10

RESULT 3  
178541  
gonadoliberin precursor - rhesus macaque (fragment)  
N:Alternate names: luteinizing hormone releasing hormone  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: I78541  
R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.  
Neuroendocrinology 60, 346-359, 1994  
A:Title: Developmental expression of the genes encoding transforming growth factor alpha  
A:Reference number: I58134; MUID:95124501  
A:Accession: I78541  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-67 <RSS>  
A:Cross-references: GB:S75318; NID:g912831; PIDN:AA33096.1; PID:g912832  
C:Superfamily: gonadoliberin

Query Match 55.8%; Score 58; DB 2; Length 67;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 7 HWSYGLRPG 15

RESULT 4  
151423  
gonadoliberin precursor - African clawed frog  
N:Alternate names: luteinizing hormone releasing hormone  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C:Accession: I51423  
R:Hayes, W.P.; Wray, S.; Batley, J.F.  
Endocrinology 134, 1835-1845, 1994  
A:Title: The frog GnRH-1 gene has a mammalian-like expression pattern and conserved domain  
A:Reference number: I51423; MUID:94185563  
A:Accession: I51423  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-89 <HAY>  
A:Cross-references: GB:I28040; NID:g496291; PIDN:AAA9728.1; PID:g496292  
C:Genetics:  
A:Gene: GnRH-1  
C:Superfamily: gonadoliberin

Query Match 55.8%; Score 58; DB 2; Length 89;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10

DB 25 HWSYGLRPG 33

RESULT 5  
RHMSG  
gonadoliberin precursor - mouse  
N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)  
C:Contains: gonadoliberin; gonadoliberin-associated protein (GAP)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1993 #sequence\_revision 18-Mar-1997 #text\_change 18-Jun-1999  
C:Accession: A47578  
R:Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli  
Science 234, 1366-1371, 1986  
A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible for the hypogonadism in the mouse  
A:Reference number: A47578; MUID:87069928  
A:Accession: A47578  
A:Molecule type: DNA  
A:Residues: 1-90 <MAS>  
A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387475  
C:Genetics:  
A:Insertions: 45/3; 77/3  
C:Function:  
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitrop  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:22-31/Product: gonadoliberin #status predicted <GLS>  
F:25-90/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 55.8%; Score 58; DB 1; Length 90;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 23 HWSYGLRPG 31

RESULT 6  
RHMSG  
gonadoliberin precursor - human  
N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)  
C:Contains: gonadoliberin-associated protein (GAP); progadoliberin  
C:Species: Homo sapiens (man)  
C:Date: 17-Mar-1987 #sequence\_revision 21-Jul-1995 #text\_change 18-Jun-1999  
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718  
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.  
Nucleic Acids Res. 17, 6403-6404, 1989  
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone gene  
A:Reference number: S05308; MUID:8936682  
A:Accession: S05308  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-92 <HAY>  
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone  
A:Reference number: A94090; MUID:86094338  
A:Accession: A26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749  
R:Seeburg, P.H.; Adelman, J.P.  
Nature 311, 666-668, 1984  
A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing hormone  
A:Reference number: A93342; MUID:85012739

A:Accession: A93342  
A:Molecule type: mRNA  
A:Residues: 1-15, 'S', 17-92 <SEE>  
A:Cross-references: GB:X01059; NID:934356; PIDN:CAA2526.1; PID:934357  
A:Experimental source: placenta  
R:Tan, L.; Rousseau, P.  
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982  
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in  
A:Reference number: A90108; MUID:83126573  
A:Accession: A90108  
A:Molecule type: protein  
A:Residues: 24-33 <TAN>  
A:Experimental source: placental trophoblasts  
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baummeister, W.; Amsterda  
FEBS Lett. 346, 203-206, 1994  
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th  
A:Reference number: S45718; MUID:94283597  
A:Contents: annotation; degradation pathway of synthetic hormone  
C:Genetics:  
A:Gene: GDB:GNRH; LHRH; GRH  
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760  
A:Map position: 8p21-8p11.2  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
F:1-23/Domains: signal sequence #status predicted <SIG>  
F:24-92/Product: progadoliberin #status predicted <PGN>  
F:24-33/Product: gonadoliberin #status experimental <MAT>  
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 55.8%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.49;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 25 HWSYGLRPG 33

RESULT 7  
RHRTG  
gonadoliberin precursor - rat  
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormo  
N:Contents: gonadoliberin; prolactin release-inhibiting factor  
C:Species: Natus norvegicus (Norway rat)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1999  
C:Accession: A40147; B26173; A48410  
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.  
Mol. Endocrinol. 3, 1257-1262, 1989  
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex  
A:Reference number: A40147; MUID:89384661  
A:Accession: A40147  
A:Molecule type: DNA  
A:Residues: 1-92 <BON>  
A:Cross-references: GB:M31670; NID:9204447; PIDN:AAA41264.1; PID:9204448  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot  
A:Reference number: A94090; MUID:86094338  
A:Accession: B26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12579; NID:9204445; PIDN:AAA41263.1; PID:9204446  
R:Waller, C.C.; Marchetti, B.; Leboeuf, R.D.; Blalock, J.E.  
Cell. Mol. Neurobiol. 12, 447-454, 1992  
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone  
A:Reference number: A48410; MUID:93105480

A:Accession: A48410  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-92 <MAT>  
A:Cross-references: GB:S50870; NID:9262059; PIDN:AB24572.1; PID:9262060  
A:Experimental source: thymus  
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)  
C:Genetics:  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: stimulates pituitary secretion of luteotropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activi  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
F:1-23/Domains: signal sequence #status predicted <SIG>  
F:24-92/Product: progadoliberin #status predicted <PGN>  
F:24-33/Product: gonadoliberin #status predicted <Gln>  
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 55.8%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.49;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 25 HWSYGLRPG 33

RESULT 8  
RNAQI  
gonadoliberin I - American alligator  
N:Alternate names: gonadotropin-releasing hormone I  
C:Species: Alligator mississippiensis (American alligator)  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Mar-1997  
C:Accession: A60066  
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan  
Regul. Pept. 33, 105-116, 1991  
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains  
A:Reference number: A60066; MUID:91352338  
A:Accession: A60066  
A:Molecule type: protein  
A:Residues: 1-10 <LOV>  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 51.9%; Score 54; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.18;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 2 HWSYGLRPG 10

RESULT 9  
150644  
gonadoliberin I precursor - chicken  
N:Alternate names: gonadotropin-releasing hormone I  
C:Species: Gallus gallus (chicken)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999  
C:Accession: 150644; S33507  
R:Dunn, T.C.; Chen, Y.; Hoek, C.; Sharp, P.J.; Sang, H.M.  
J. Mol. Endocrinol. 11, 19-29, 1993  
A:Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene  
A:Reference number: 150644; MUID:94059355  
A:Accession: 150644  
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-92 <DU2>  
A:Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA9246.1; PID:g311612  
C:Genetics: 47/3; 79/3  
C:Superfamily: gonadolibertin

Query Match 51.9%; Score 54; DB 2; Length 92;  
Best Local Similarity 88.9%; Pred. No. 1.6;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
| | | | | | | |  
Db 25 HWSYGLQPG 33

RESULT 10  
150739  
gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)

C:Species: Haplochromis burtoni  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: 150739  
R:White, S.A.; Kastan, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.  
Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995  
A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles  
A:Reference number: 150739; MUID:95396797  
A:Accession: 150739  
A:Status: preliminary; translated from GB/EMBL/DBR  
A:Molecule type: mRNA  
A:Residues: 1-98 <WHI>  
A:Cross-references: EMBL:U31865; NID:g905398; PIDN:AAC59691.1; PID:g905399  
C:Superfamily: gonadolibertin

Query Match 50.0%; Score 52; DB 2; Length 98;  
Best Local Similarity 88.9%; Pred. No. 3.1;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
| | | | | | | |  
Db 24 HWSYGLSPG 32

RESULT 11  
A30996  
orf-X protein - simian immunodeficiency virus (isolate Mne) (fragments)

N:Alternate names: VPX protein  
C:Species: simian immunodeficiency virus, SIV  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 08-Apr-1994  
C:Accession: A30996  
R:Henderson, L.E.; Sowder, R.C.; Copeland, T.D.; Benveniste, R.E.; Oroszian, S.  
Science 241, 199-201, 1988  
A:Title: Isolation and characterization of a novel protein (X-ORF product) from SIV and  
A:Reference number: A30996; MUID:88264422  
A:Accession: A30996  
A:Molecule type: protein  
A:Residues: 1-106 <HEN>  
C:Genetics:  
A:Gene: VPX  
C:Superfamily: AIDS vpu protein

Query Match 50.0%; Score 52; DB 2; Length 106;  
Best Local Similarity 60.0%; Pred. No. 3.4;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPP 16  
| : | | | | | | | |  
Db 88 HGAAGWRGPPPP 102

RESULT 12

ASLX3  
vpu protein - simian immunodeficiency virus (macaque isolate)

N:Alternate names: orf-X protein  
C:Species: simian immunodeficiency virus, SIV  
C>Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 16-Jul-1999  
C:Accession: D28887  
R:Chakrabarti, L.; Guyader, M.; Allison, M.; Daniel, M.D.; Desrosiers, R.C.; Tiollais, P.  
Nature 328, 543-547, 1987  
A:Title: Sequence of simian immunodeficiency virus from macaque and its relationship  
A:Reference number: A28887; MUID:87287230  
A:Accession: D28887  
A:Molecule type: DNA  
A:Residues: 1-112 <CHA>  
A:Cross-references: GB:Y00277; GB:M16403; NID:g61730; PIDN:CAA68382.1; PID:g61734  
C:Genetics:  
A:Gene: vpu  
C:Superfamily: AIDS vpu protein  
C:Keywords: AIDS; immunodeficiency

Query Match 50.0%; Score 52; DB 1; Length 112;  
Best Local Similarity 60.0%; Pred. No. 3.5;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPP 16  
| : | | | | | | | |  
Db 94 HGAAGWRGPPPP 108

RESULT 13

S03066  
gene X protein - human T-cell lymphotropic virus type 4  
C:Species: human T-cell lymphotropic virus type 4, HTLV-4  
C>Date: 21-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999  
C:Accession: S03066  
R:Hahn, B.H.; Kong, L.I.; Lee, S.W.; Kumar, P.; Taylor, M.E.; Arya, S.K.; Shaw, G.M.  
Nature 300, 184-186, 1987  
A:Title: Relation of HTLV-4 to simian and human immunodeficiency-associated viruses.  
A:Reference number: S03065  
A:Accession: S03066  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-112 <HAH>  
A:Cross-references: EMBL:X06391; NID:g61580; PIDN:CAA29688.1; PID:g61581  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1987  
C:Superfamily: AIDS vpu protein

Query Match 50.0%; Score 52; DB 2; Length 112;  
Best Local Similarity 60.0%; Pred. No. 3.5;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPP 16  
| : | | | | | | | |  
Db 94 HGAAGWRGPPPP 108

RESULT 14

ASLST  
vpu protein - simian immunodeficiency virus SIVagm (type 3, isolate SIVLV-3agm)  
N:Alternate names: orf-X protein  
C:Species: simian immunodeficiency virus SIVagm  
C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 31-Jan-1997  
C:Accession: D26737  
R:Hirsch, V.  
Submitted to GenBank, June 1987

A:Reference number: A26737  
A:Accession: D26737  
A:Molecule type: DNA  
A:Residues: 1-112 <HIR>  
C:Genetics:  
A:Gene: vpu  
C:Superfamily: AIDS vpu protein

C:Keywords: AIDS; immunodeficiency

Query Match 49.0%; Score 51; DB 1; Length 112;  
Best Local Similarity 60.0%; Pred. No. 4.8;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPP 16  
| : | | | | | | |  
Db 94 HRAGGWRPGPPPPP 108

RESULT 15

T46289

hypothetical protein DKFZp434A10.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000

C:Accession: T46289

R:Duisterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000

A:Reference number: 223035

A:Accession: T46289

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-862 <AAA>

A:Cross-references: EMBL:AL137579

A:Experimental source: adult testis; clone DKFZp434A10.1

C:Genetics:

A:Note: DKFZp434A10.1

Query Match

Best Local Similarity 49.0%; Score 51; DB 2; Length 862;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 LRPSSPPPP 16  
| | | : | | | |  
Db 458 LRPSSPPPP 467

Search completed: March 2, 2001, 10:55:40  
Job time: 260 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:36 ; Search time 47.48 Seconds  
(without alignments)  
6.429 Million cell updates/sec

Title: US-09-306-689-3

Perfect score: 108

Sequence: 1 CPPPSEHWSYGLRPG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents, AA: \*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/Backfillies1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	102	94.4	17 1	US-08-188-223-5
2	102	94.4	17 3	US-08-968-466-5
3	82.5	76.4	16 1	US-08-188-223-7
4	82.5	76.4	16 3	US-08-968-466-7
5	81	75.0	34 3	US-09-100-414B-82
6	81	75.0	35 3	US-09-100-414B-80
7	81	75.0	35 3	US-09-100-414B-81
8	71	65.7	49 1	US-08-387-156-4
9	71	65.7	49 2	US-08-694-865-4
10	71	65.7	49 2	US-08-878-748-4
11	71	65.7	49 2	US-09-124-491-4
12	71	65.7	544 1	US-08-387-156-10
13	71	65.7	544 2	US-08-694-865-10
14	71	65.7	544 2	US-08-878-748-10
15	71	65.7	544 2	US-09-124-491-10
16	71	65.7	699 2	US-08-694-865-16
17	71	65.7	699 3	US-09-124-491-16
18	71	65.7	977 1	US-08-387-156-8
19	71	65.7	977 2	US-08-694-865-8
20	71	65.7	977 2	US-08-878-748-8
21	71	65.7	977 2	US-09-124-491-8
22	68	63.0	14 1	US-07-690-983D-24
23	68	63.0	24 1	US-07-690-983D-43
24	68	63.0	44 1	US-07-690-983D-45
25	68	63.0	84 1	US-07-690-983D-47
26	66	61.1	90 1	US-08-341-219-19
27	66	61.1	17 1	US-07-690-983D-16
28	66	61.1	28 1	US-08-446-692-10

29	66	61.1	28 2	US-08-488-351A-10	Sequence 10, Appl
30	65	60.2	14 1	US-07-690-983D-30	Sequence 30, Appl
31	64	59.3	16 1	US-08-453-588-26	Sequence 26, Appl
32	64	59.3	16 3	US-08-521-079-26	Sequence 6, Appl
33	64	59.3	52 3	US-08-458-814-6	Sequence 6, Appl
34	64	59.3	55 3	US-08-458-814-7	Sequence 7, Appl
35	64	59.3	936 1	US-08-455-970A-12	Sequence 12, Appl
36	63	58.3	10 1	US-07-714-540-9	Sequence 9, Appl
37	63	58.3	10 1	US-07-690-983D-2	Sequence 2, Appl
38	63	58.3	10 1	US-07-690-983D-32	Sequence 32, Appl
39	63	58.3	10 1	US-08-343-883-1	Sequence 1, Appl
40	63	58.3	10 1	US-08-000-931-5	Sequence 5, Appl
41	63	58.3	10 1	US-08-428-488-22	Sequence 22, Appl
42	63	58.3	10 1	US-08-341-219-11	Sequence 11, Appl
43	63	58.3	10 1	US-08-453-588-22	Sequence 22, Appl
44	63	58.3	10 1	US-08-591-917-1	Sequence 1, Appl
45	63	58.3	10 5	5492893-1	Patent No. 5492893

#### ALIGNMENTS

RESULT 1  
US-08-188-223-5  
Sequence 5, Application US/08188223  
Patent No. 5688506  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Scibienski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,223  
FILING DATE: 27-JAN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: YES  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..7  
OTHER INFORMATION: /note= "spacer"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 8..17  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 17

OTHER INFORMATION: /label=glyNH2  
OTHER INFORMATION: /note="glycinamide"  
US-08-188-223-5

Query Match 94.4%; Score 102; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CPPPSEHWSYGLRP 16  
Db 1 CPPPSEHWSYGLRP 16

RESULT 2  
US-08-968-466-5  
Sequence 5, Application US/08968466  
Patent No. 6132720  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Scibienski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968,466  
FILING DATE: 27-JAN-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEO ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: YES  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..7  
OTHER INFORMATION: /note="spacer"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 8..17  
OTHER INFORMATION: /note="immunomimic"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 17  
OTHER INFORMATION: /label=glyNH2  
OTHER INFORMATION: /note="glycinamide"  
US-08-968-466-5

Query Match 94.4%; Score 102; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPPSEHWSYGLRP 16  
Db 1 CPPPSEHWSYGLRP 16

RESULT 3  
US-08-188-223-7  
Sequence 7, Application US/08188223  
Patent No. 5688506  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Scibienski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,223  
FILING DATE: 27-JAN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEO ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: YES  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..6  
OTHER INFORMATION: /note="spacer"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 7..16  
OTHER INFORMATION: /note="immunomimic"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 16  
OTHER INFORMATION: /label=glyNH2  
OTHER INFORMATION: /note="glycinamide"  
US-08-188-223-7

Query Match 76.4%; Score 82.5; DB 1; Length 16;  
Best Local Similarity 87.5%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CPPPSEHWSYGLRP 16  
Db 1 CPPPP-REHWSYGLRP 15

RESULT 4  
US-08-968-466-7

Sequence 7, Application US/08968466  
Patent No. 6132720  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Scibienski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968,466  
FILING DATE: 27-JAN-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-354-8113  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..6  
OTHER INFORMATION: /note="spacer"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 7..16  
OTHER INFORMATION: /note="immunomimic"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 16  
OTHER INFORMATION: /label="GLYNH2"  
OTHER INFORMATION: /note="glycinamide"  
US-08-968-466-7

Query Match 76.4%; Score 82.5; DB 3; Length 16;  
Best Local Similarity 87.5%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CPPPSSEHWSYGLRP 16  
DB 1 CPPPP-REHWSYGLRP 15

RESULT 5  
US-09-100-414B-82  
Sequence 82, Application US/09100414B  
Patent No. 6025468  
GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
TITLE OF INVENTION: IMMUNOGENS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,414B  
FILING DATE: 20-JUNE-1998  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 80:

Query Match 75.0%; Score 81; DB 3; Length 34;  
Best Local Similarity 81.2%; Pred. No. 4.8e-05;  
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PPPPSSEHWSYGLRPG 17  
DB 19 PPXPPEHWSYGLRPG 34

RESULT 6  
US-09-100-414B-80  
Sequence 80, Application US/09100414B  
Patent No. 6025468  
GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
TITLE OF INVENTION: IMMUNOGENS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,414B  
FILING DATE: 20-JUNE-1998  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 80:



APPLICATION NUMBER: US/08/694,865  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-865-4

Query Match 65.7%; Score 71; DB 2; Length 49;  
Best Local Similarity 80.0%; Pred. No. 0.0016;  
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 PPPSEHWSYGLRPG 17  
1 ||:|||||  
DB 22 PGSSSQHWSYGLRPG 36

RESULT 10  
US-08-878-748-4  
Sequence 4, Application US/08878748  
Patent No. 5969126  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 633 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/878,748  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-878-748-4

Query Match 65.7%; Score 71; DB 2; Length 49;  
Best Local Similarity 80.0%; Pred. No. 0.0016;  
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 PPPSEHWSYGLRPG 17  
1 ||:|||||  
DB 22 PGSSSQHWSYGLRPG 36

RESULT 11  
US-09-124-491-4  
Sequence 4, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-4

Query Match 65.7%; Score 71; DB 3; Length 49;  
Best Local Similarity 80.0%; Pred. No. 0.0016;  
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 PPPSEHWSYGLRPG 17  
1 ||:|||||  
DB 22 PGSSSQHWSYGLRPG 36



REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-878-748-10

Query Match 65.7%; Score 71; DB 2; Length 544;  
Best Local Similarity 80.0%; Pred. No. 0.02;  
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPSSEHWSYGLRPG 17  
1 ||:|||||  
DB 515 PGSSSQHWSYGLRPG 529

RESULT 15  
US-09-124-491-10  
Sequence 10, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANNIS, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-10

Query Match 65.7%; Score 71; DB 3; Length 544;  
Best Local Similarity 80.0%; Pred. No. 0.02;  
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPSSEHWSYGLRPG 17  
1 ||:|||||  
DB 515 PGSSSQHWSYGLRPG 529

Search completed: March 2, 2001, 10:54:36  
Job time: 601 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:00:45 ; Search time 97.13 Seconds  
(without alignments)  
20.514 Million cell updates/sec

Title: US-09-306-689-3  
Perfect score: 108  
Sequence: 1 CPEPSEHMSYGLRPG 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues  
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_15:\*  
2: sp.archaea:\*  
3: sp.bacteria:\*  
4: sp.fungi:\*  
5: sp.human:\*  
6: sp.invertebrate:\*  
7: sp.mammal:\*  
8: sp.mhc:\*  
9: sp.organelle:\*  
10: sp.phage:\*  
11: sp.plant:\*  
12: sp.todent:\*  
13: sp.virus:\*  
14: sp.vertebrate:\*  
15: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	55.6	91	13	O9PRH0
2	54	50.0	87	13	O9YI26
3	54	50.0	95	13	O73812
4	54	50.0	99	13	O91A10
5	54	50.0	114	6	O97655
6	52	48.1	90	13	O91A02
7	52	48.1	110	6	O97686
8	52	48.1	148	4	O9WZ81
9	52	48.1	2357	5	O9UIM8
10	51	47.2	33	13	O9W760
11	51	47.2	33	13	O9PT34
12	51	47.2	82	13	O92094
13	51	47.2	82	13	O9W7G1
14	51	47.2	82	13	O91800
15	51	47.2	82	13	O918P9
16	51	47.2	88	13	O9PSY9
17	51	47.2	107	6	O91A09
18	51	47.2	107	6	O9YTS13
19	51	47.2	327	4	O9UHN2

20	51	47.2	327	4	O9UBV7	O9UBV7 homo sapien
21	50.5	46.8	2942	12	O9WB76	O9WB76 gp virus c
22	50	46.3	86	13	O42471	O42471 carassius a
23	50	46.3	86	13	O9PT25	O9PT25 oncorhynch
24	50	46.3	230	5	O9W3V7	O9W3V7 dirosophila
25	50	46.3	497	5	O18662	O18662 caenorhabd
26	49.5	45.8	89	2	O9ZNI3	O9ZNI3 pseudomonas
27	49.5	45.8	437	6	O9RS25	O9RS25 bos taurus
28	49.5	45.8	473	5	O9YOS9	O9YOS9 dirosophila
29	49	45.4	816	5	O97214	O97214 leishmania
30	49	45.4	880	5	O9W575	O9W575 dirosophila
31	48	44.4	348	5	O9VSB8	O9VSB8 dirosophila
32	47.5	44.0	305	5	O15844	O15844 leishmania
33	47.5	44.0	1044	4	O9P206	O9P206 homo sapien
34	47	43.5	75	10	O01948	O01948 lycopersico
35	47	43.5	190	2	O86520	O86520 streptomyce
36	47	43.5	226	2	O69910	O69910 streptomyce
37	47	43.5	322	10	O09084	O09084 lycopersico
38	47	43.5	325	10	O81467	O81467 arabidopsis
39	47	43.5	327	5	O9TYP2	O9TYP2 caenorhabd
40	47	43.5	612	4	O75802	O75802 homo sapien
41	47	43.5	788	5	O96416	O96416 dirosophila
42	46.5	43.1	267	4	O13261	O13261 homo sapien
43	46	42.6	164	2	O9PCB4	O9PCB4 xyella fas
44	46	42.6	201	10	O9XIN6	O9XIN6 arabidopsis
45	46	42.6	252	10	O9SD40	O9SD40 arabidopsis

## ALIGNMENTS

RESULT	ID	Query Match	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
1	O9PRH0	PRELIMINARY:	PRT:	91	AA.					
AC	O9PRH0:									
DT	01-MAY-2000 (TREMBLrel. 13, Created)									
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)									
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)									
DE	PREPRO-MGNRH PRECURSOR.									
OS	Anguilla japonica (Japanese eel).									
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;									
OC	Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;									
OX	NCBI_TaxID=7937;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	TISSUE=BRAIN;									
RA	Okubo K., Suetake H., Aida K.;									
RT	"Expression of two gonadotropin-releasing hormone (GNRH) precursor									
RT	genes in various tissues of the Japanese eel and evolution of GNRH.";									
RL	Zool. Sci. 16:471-478(1999).									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RA	Okubo K., Suetake H., Aida K.;									
RT	"A splicing variant for the prepro-mammalian gonadotropin-releasing									
RT	hormone (prepro-mGNRH) mRNA is present in the brain and various									
RL	peripheral tissues of the Japanese eel.";									
RL	Zool. Sci. 16:645-651(1999).									
DR	EMBL; AB026989; BAA82608.1;									
DR	EMBL; AB026991; BAA83597.1;									
DR	INTERPRO: IPR02012;									
DR	PFAM; PF00446; GNRH; 1.									
DR	PROSITE; PS00473; GNRH; 1.									
KW	SIGNAL.									
FT	SIGNAL.									
FT	CHAIN									
FT	CHAIN									
SO	SEQUENCE									

Query Match 55.6%; Score 60; DB 13; Length 91;  
Best Local Similarity 90.0%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 EHWSTGLRPG 17  
:|||||||  
DB 23 QHWSTGLRPG 32

RESULT 2  
Q9Y126 PRELIMINARY; PRT; 87 AA.

AC 09Y126: 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)  
DE (FRAGMENT).  
OS Sparus aurata (Gilthead sea bream).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
OC Sparidae; Sparus.  
OX NCBI\_TaxID=8175;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Ovary;  
RA Nabissi M.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
FOLLICLE-STIMULATING HORMONES.  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
DR EMBL; AF046801; AAD02427.1; -.  
DR INTERPRO; IPR002012; -.  
DR PFAM; PF00446; GNRH; 1.  
DR PROSITE; PS00473; GNRH; 1.  
KW Hormone; Amidation.  
FT NON\_TER 1  
FT 87  
FT SEQUENCE 87 AA; 9871 MW; 0D246353D96782A CRC64;

Query Match 50.0%; Score 54; DB 13; Length 87;  
Best Local Similarity 80.0%; Pred. No. 0.75;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 EHWSTGLRPG 17  
:|||||||  
DB 21 QHWSTGLRPG 30

RESULT 3  
Q73812 PRELIMINARY; PRT; 95 AA.

AC 073812: 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).  
OS Morone saxatilis (Striped bass).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
OC Moronidae; Morone.  
OX NCBI\_TaxID=34816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Chow M.M., Kight K.E., Gotthelf Y., Alok D., Zohar Y.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
FOLLICLE-STIMULATING HORMONES.  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
DR EMBL; AF056314; AAD03817.1; -.  
DR INTERPRO; IPR002012; -.  
DR PFAM; PF00446; GNRH; 1.  
DR PROSITE; PS00473; GNRH; 1.

DR PRODOM; PD005581; -; 1.  
KW Hormone; Amidation.  
SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match 50.0%; Score 54; DB 13; Length 95;  
Best Local Similarity 80.0%; Pred. No. 0.81;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 EHWSTGLRPG 17  
:|||||||  
DB 23 QHWSTGLRPG 32

RESULT 4  
Q9IA10 PRELIMINARY; PRT; 99 AA.

AC 09IA10: 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE GONADOTROPIN-RELEASING HORMONE SEABREAM ISOFORM.  
OS Dicentrarchus labrax (European sea bass).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
OC Moronidae; Dicentrarchus.  
OX NCBI\_TaxID=13489;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanuy S.,  
RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;  
RL "Differential expression of three different prepro-GNRH  
(gonadotrophin-releasing hormone) messengers in the brain of the  
European sea bass (Dicentrarchus labrax)."  
RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA Zmora N., Zohar Y., Elizur A.;  
RL "3 GNRH form in the seabass Dicentrarchus labrax."  
RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF224279; AAF62898.1; -.  
SQ SEQUENCE 99 AA; 10758 MW; EC8AEEC93CC02904 CRC64;

Query Match 50.0%; Score 54; DB 13; Length 99;  
Best Local Similarity 80.0%; Pred. No. 0.85;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 EHWSTGLRPG 17  
:|||||||  
DB 27 QHWSTGLRPG 36

RESULT 5  
Q97655 PRELIMINARY; PRT; 114 AA.

AC 097655: 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)  
DE PRECURSOR.  
GN GNRH2.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.

RA White R.B., Urbanski H.F., Fernald R.D.;  
 RT "A second gene for gonadotropin-releasing hormone is expressed in the  
 RL rhesus macaque (Abstract #632.18)";  
 CC Abstr. Soc. Neurosci. 24:1609-1609(1998).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 DR EMBL: AF097356; AAD09106.1; -;  
 DR INTERPRO: IPR002012; -;  
 DR PIRAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 DR PRODOM: PD029787; -; 1.  
 KW Hormone; Amidation.  
 SQ SEQUENCE 114 AA; 12533 MW; 8B70D690D5BD5103 CRC64;

Query Match 50.0%; Score 54; DB 6; Length 114;  
 Best Local Similarity 50.0%; Pred. No. 0.97;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 PPPPSSEHWSYGLRPG 17  
 Db 19 PGPESEAHWSHGWPY 34

RESULT 6  
 O91AU2 PRELIMINARY; PRT; 90 AA.  
 AC O91AU2;  
 DT 01-OCT-2000 (TREMBLREL. 15, Created)  
 DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)  
 DE GONADOTROPIN-RELEASING HORMONE.  
 OS Rana dybowskii (Trog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.  
 OX NCBI\_TaxId=71582;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RA Yoo M.S., Kang H.M., Choi H.S., Chun S.Y., Troskie B., Millar R.P.,  
 RA Kwon H.B.;  
 RT "Molecular Cloning, Distribution and Pharmacological Characterization  
 RT of a Novel Gonadotropin-releasing Hormone([Trp8]GNRH) in Frog Brain";  
 RL Mol. Cell. Endocrinol. 0:0-0(2000).  
 DR EMBL: AF139111; AAF44343.1; -;  
 SQ SEQUENCE 90 AA; 10368 MW; C3D573E78B52ABFA CRC64;

Query Match 48.1%; Score 52; DB 13; Length 90;  
 Best Local Similarity 80.0%; Pred. No. 1.5;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 8 EHWSYGLRPG 17  
 Db 25 QHWSYGLMPC 34

RESULT 7  
 O97686 PRELIMINARY; PRT; 110 AA.  
 AC O97686;  
 DT 01-MAY-1999 (TREMBLREL. 10, Created)  
 DT 01-MAY-1999 (TREMBLREL. 10, Last sequence update)  
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).  
 GN GNRH2.  
 OS Suncus murinus (House shrew).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Insectivora; Soricidae; Suncus.  
 OX NCBI\_TaxId=9378;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN.  
 RA White R.B., Kasten T.L., White S.A., Rissman E.F., Fernald R.D.;  
 RT "GNRH-II cDNA expression in the musk shrew";  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 DR EMBL: AF107315; AAD09114.1; -;  
 DR INTERPRO: IPR002012; -;  
 DR PIRAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 DR PRODOM: PD029787; -; 1.  
 KW Hormone; Amidation.  
 SQ SEQUENCE 110 AA; 12120 MW; A8986905FB83D9DB CRC64;

Query Match 48.1%; Score 52; DB 6; Length 110;  
 Best Local Similarity 50.0%; Pred. No. 1.8;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 PPPPSSEHWSYGLRPG 17  
 Db 21 PGPLKAQHSWGWPY 36

RESULT 8  
 O9N281 PRELIMINARY; PRT; 148 AA.  
 AC O9N281;  
 DT 01-OCT-2000 (TREMBLREL. 15, Created)  
 DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)  
 DE UNCHARACTERIZED BONE MARROW PROTEIN BM041.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BONE MARROW;  
 RA Zhao M., Gu J., Li N., Peng Y., Han Z., Chen Z.;  
 RT "A novel gene expressed in human bone marrow";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF217517; AAF67628.1; -;  
 SQ SEQUENCE 148 AA; 15385 MW; 56BFE6A15935A2E CRC64;

Query Match 48.1%; Score 52; DB 4; Length 148;  
 Best Local Similarity 47.1%; Pred. No. 2.4;  
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 CPPPSSEHWSYGLRPG 17  
 Db 41 CPPPGAPHGNGPAFP 57

RESULT 9  
 O90IM8 PRELIMINARY; PRT; 2357 AA.  
 AC O90IM8;  
 DT 01-MAY-2000 (TREMBLREL. 13, Created)  
 DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)  
 DE CLASS VII UNCONVENTIONAL MYOSIN.  
 GN MYO1.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Dictyostellida; Dictyostellum.  
 OX NCBI\_TaxId=44889;  
 RN [1]  
 RP SEQUENCE OF 254-349 FROM N.A.  
 RX MEDLINE=95023928; PubMed=7937787;  
 RA Titus M.A., Kuspa A., Loomis W.F.;  
 RT "Discovery of myosin genes by physical mapping in Dictyostelium";

RL Proc. Natl. Acad. Sci. U.S.A. 91:9446-9450(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Titus M.A.;  
 RT "A class VII unconventional myosin is required for phagocytosis."  
 RL Curr. Biol. 9:1297-1303(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Titus M.A., Loomis W.F.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Titus M.A.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR HSSP; P08799; 1MND.  
 DR INTERPRO: IPR000048; -  
 DR INTERPRO: IPR000159; -  
 DR INTERPRO: IPR000299; -  
 DR INTERPRO: IPR000857; -  
 DR INTERPRO: IPR001452; -  
 DR INTERPRO: IPR001609; -  
 DR PFAM: PF000018; SH3; 1.  
 DR PFAM: PF00063; myosin\_head; 1.  
 DR PFAM: PF00612; IQ; 1.  
 DR PFAM: PF00784; MYTH4; 2.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR PROSITE: PS50002; SH3; 1.  
 DR PROSITE: PS50057; BAND; 41; 3; 2.  
 SO SEQUENCE 2357 AA; 268281 MW; FC667F12F207B0C2 CRC64;

Query Match 48.1%; Score 52; DB 5; Length 2357;  
 Best Local Similarity 60.0%; Pred. No. 34;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 PPPSSSEHWSYGLRP 16  
 DB 1027 PPPSSSSMDGLPP 1041

RESULT 10  
 O9W7G0 PRELIMINARY; PRT; 33 AA.  
 AC O9W7G0;  
 DT 01-NOV-1999 (TREMBLREL. 12, Created)  
 DT 01-NOV-1999 (TREMBLREL. 12, Last sequence update)  
 DT 01-MAY-2000 (TREMBLREL. 13, Last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN) (FRAGMENT).  
 GN GNRH2.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OC NCBI\_TaxID=8022;  
 OX [1]  
 RN RP SEQUENCE FROM N.A.  
 RX MEDLINE-99312119; PubMed-10385393;  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RT "Regulation and expression of gonadotropin-releasing hormone gene differs in brain and gonads in rainbow trout."  
 RL Endocrinology 140:3012-3024(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
 DR EMBL: AF110993; AAD3463.1; -  
 DR INTERPRO: IPR002012; -  
 DR PFAM: PF00446; GNRH; 1.

DR PROSITE: PS00473; GNRH; 1.  
 KW Hormone; Amidation.  
 FT NON\_TER 33  
 SO SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;

Query Match 47.2%; Score 51; DB 13; Length 33;  
 Best Local Similarity 72.7%; Pred. No. 0.8;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 DB 23 SQHWSYGLWLP 33

RESULT 11  
 O9PT34 PRELIMINARY; PRT; 33 AA.  
 AC O9PT34;  
 DT 01-MAY-2000 (TREMBLREL. 13, Created)  
 DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLREL. 14, Last annotation update)  
 DE GONADOTROPIN-RELEASING HORMONE (FRAGMENT).  
 GN GNRH1.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OC NCBI\_TaxID=8022;  
 OX [1]  
 RN RP SEQUENCE FROM N.A.  
 RX MEDLINE-99312119; PubMed-10385393;  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RT "Regulation and expression of gonadotropin-releasing hormone gene differs in brain and gonads in rainbow trout."  
 RL Endocrinology 140:3012-3024(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF110533; AAD3461.1; -  
 DR INTERPRO: IPR002012; -  
 DR INTERPRO: IPR002047; -  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00256; AKH; UNKNOWN\_1.  
 DR PROSITE: PS00473; GNRH; 1.  
 FT NON\_TER 33  
 SO SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;

Query Match 47.2%; Score 51; DB 13; Length 33;  
 Best Local Similarity 72.7%; Pred. No. 0.8;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 DB 23 SQHWSYGLWLP 33

RESULT 12  
 O92094 PRELIMINARY; PRT; 82 AA.  
 AC O92094;  
 DT 01-NOV-1996 (TREMBLREL. 01, Created)  
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)  
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN) (FRAGMENT).  
 GN PREPRO-GNRH-1.  
 OS Oncorhynchus nerka (Sockeye salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

OX NCBI\_TaxID=8023;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NIKKO; TISSUE=BRAIN;  
 RA MEDLINE=96020547; PubMed=8546809;  
 RA Ashihara M., Suzuki M., Kubokawa K., Yoshitura Y., Kobayashi M.,  
 RA Urano A., Aida K.;  
 RT "Two differing precursor genes for the salmon-type gonadotropin-  
 RT releasing hormone exist in salmonids.";  
 RL J. Mol. Endocrinol. 15:1-9(1995).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 DR EMBL: D31868; BAA06666.1; -  
 DR INTERPRO: IPR002012; -  
 DR INTERPRO: IPR002047; -  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00256; AKH; UNKNOWN\_1.  
 DR PRODOM: PD006549; -; 1.  
 KW Signal; Hormone; Amidation.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 33 GNRH-ASSOCIATED PEPTIDE.  
 FT CHAIN 37 82 GNRH-ASSOCIATED PEPTIDE.  
 SQ SEQUENCE 82 AA; 9126 MW; C64044EA521B2B8B CRC64;

Query Match 47.2%; Score 51; DB 13; Length 82;  
 Best Local Similarity 72.7%; Pred. No. 1.9;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 DB 23 SQHWSYGLRPG 33

RESULT 13  
 ID 09W7G1 PRELIMINARY; PRT: 82 AA.  
 AC 09W7G1;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).  
 GN GNRH1.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99312119; PubMed=10385393;  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RT "Regulation and expression of gonadotropin-releasing hormone gene  
 RT differs in brain and gonads in rainbow trout.";  
 RL Endocrinology 140:3012-3024(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 DR EMBL: AF110992; AAD3462.1; -  
 DR INTERPRO: IPR002012; -  
 DR INTERPRO: IPR002047; -  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00256; AKH; UNKNOWN\_1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KW Hormone; Amidation.  
 SQ SEQUENCE 82 AA; 9232 MW; 7595B4FCC65FDFD6 CRC64;

Query Match 47.2%; Score 51; DB 13; Length 82;  
 Best Local Similarity 72.7%; Pred. No. 1.9;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 DB 23 SQHWSYGLRPG 33

RESULT 14  
 ID 091800 PRELIMINARY; PRT: 82 AA.  
 AC 091800;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE PRO-SGNRH-1.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RA Ferriere F., Bailhache T., Jégo P.;  
 RT "Oncorhynchus mykiss sgnrh-1 cDNA from brain.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF232212; AAF91280.1; -  
 SQ SEQUENCE 82 AA; 9198 MW; 7595A0B696556A69 CRC64;

Query Match 47.2%; Score 51; DB 13; Length 82;  
 Best Local Similarity 72.7%; Pred. No. 1.9;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 DB 23 SQHWSYGLRPG 33

RESULT 15  
 ID 0918P9 PRELIMINARY; PRT: 82 AA.  
 AC 0918P9;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE PRO-SGNRH-1.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RA Ferriere F., Bailhache T., Jégo P.;  
 RT "Oncorhynchus mykiss sgnrh-1 cDNA in the brain.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF232213; AAF91281.1; -  
 SQ SEQUENCE 82 AA; 9203 MW; 8053F4F221A0F08 CRC64;

Query Match 47.2%; Score 51; DB 13; Length 82;  
 Best Local Similarity 72.7%; Pred. No. 1.9;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 DB 23 SQHWSYGLRPG 33

Fri Mar 2 12:19:30 2001

us-09-306-689-3.rspt

Page 6

Search completed: March 2, 2001, 11:00:47  
Job time: 457 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:01:19 ; Search time 32.03 Seconds  
(without alignments)  
17.140 Million cell updates/sec

Title: US-09-306-689-3  
Perfect score: 108  
Sequence: 1 CPPPSEHWSYGLRPG 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues  
Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	63.9	92	1	GON1_HUMAN
2	68	63.0	92	1	GON1_MACMU
3	68	63.0	90	1	GON1_MOUSE
4	68	63.0	91	1	GON1_PIG
5	68	63.0	92	1	GON1_RAT
6	61	56.5	89	1	GON1_XENLA
7	61	56.5	92	1	GON1_TENBLA
8	60	55.6	61	1	GON1_SHEEP
9	60	55.6	63	1	GON1_MESAU
10	57	52.8	92	1	GON1_CAVPO
11	57	52.8	92	1	GON1_CHICK
12	56	51.9	10	1	GON1_ALMT
13	56	51.9	94	1	GON1_HAPBU
14	55	50.9	80	1	GON1_CLAGA
15	54	50.0	95	1	GON1_PAGMA
16	54	50.0	95	1	GON1_SPAU
17	54	50.0	114	1	GON2_TUPGB
18	51	47.2	74	1	GON3_ONCMY
19	51	47.2	74	1	GON3_ONCMY
20	51	47.2	82	1	GON3_ONCMY
21	51	47.2	82	1	GON3_SALSA
22	51	47.2	82	1	GON3_SALSA
23	51	47.2	89	1	GON3_SATFR
24	51	47.2	90	1	GON3_PORNO
25	51	47.2	90	1	GON3_HAPBU
26	51	47.2	90	1	GON3_PAGMA
27	50	46.3	86	1	GON2_SPAU
28	50	46.3	86	1	GON2_CARAU
29	50	46.3	86	1	GON2_ONCMY
30	49.5	45.8	437	1	GON2_RUTRU
31	49	45.4	120	1	PTB2_BOVIN
32	48	44.4	906	1	GON2_HUMAN
33	47	43.5	10	1	KBF2_CHICK
					GON3_ONCMY

## ALIGNMENTS

RESULT ID	1	GON1_HUMAN	STANDARD:	PRT:	92 AA.
AC	GON1_HUMAN				
AD	P01148:				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	01-APR-1988 (Rel. 07, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I) (LUTEINIZING HORMONE RELEASING HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED PEPTIDE I].				
DE	HORMONE I] (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED PEPTIDE I].				
DE	GNRH I OR GNRH OR LHRH.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-8936682; PubMed-2671939;				
RA	Hayflick J.S., Adelman J.P., Seeburg P.H.;				
RT	"The complete nucleotide sequence of the human gonadotropin-releasing hormone gene."				
RL	Nucleic Acids Res. 17:6403-6403(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-8609438; PubMed-2867548;				
RA	Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;				
RT	"Characterization of cDNA for precursor of human luteinizing hormone releasing hormone."				
RL	Nature 311:666-668(1984).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-85012739; PubMed-6090951;				
RA	Seeburg P.H., Adelman J.P.;				
RT	"Characterization of cDNA for precursor of human luteinizing hormone releasing hormone."				
RL	Nature 311:666-668(1984).				
RN	[4]				
RP	SEQUENCE OF 24-33.				
RX	MEDLINE-83126573; PubMed-6760865;				
RA	Tan L., Rousseau P.;				
RT	"The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in the human placenta."				
RL	Biochem. Biophys. Res. Commun. 109:1061-1071(1982)				
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.				
CC	-1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS), LUTREPISE OR LUTRELEF (FERRING PHARMACEUTICALS) AND RELISORM (SEERON).				
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way				

34	47	43.5	94	1	GON3_CARAU	P51917 carcassius a
35	47	43.5	94	1	GON3_RUTRU	O92106 rutilus rut
36	47	43.5	396	1	FLO_ANTMA	P23915 antirrhinum
37	47	43.5	757	1	CIKE_HUMAN	Q14003 homo sapien
38	47	43.5	769	1	CIKE_MOUSE	O63959 mus musculu
39	47	43.5	889	1	CIKE_RAT	O01956 rattus norv
40	46.5	43.1	250	1	CPBL_CHICK	P81475 gallus gall
41	46	42.6	86	1	GON2_CLAGA	P43306 claritas gar
42	46	42.6	95	1	DAMS_HUMAN	O99013 homo sapien
43	46	42.6	551	1	CAP_SCHPO	P36621 schizosacch
44	46	42.6	756	1	PIP6_RAT	P10688 rattus norv
45	45	41.7	383	1	DEMA_HUMAN	Q08495 homo sapien

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL: X01059; CAA25526.1; -  
DR EMBL: M12578; AAA35916.1; -  
DR EMBL: X15215; CAA33285.1; -  
DR PIR: A01410; RHUG.  
DR PIR: A26173; A26173.  
DR PIR: S05308; S05308.  
DR MIM: 152760; -  
DR INTERPRO: IPR002012; -  
DR PFAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Placenta; Pharmaceutical; Signal.

FT SIGNAL 1 23  
FT CHAIN 24 92 PROGNADOLIBERIN I.  
FT PEPTIDE 24 33 GONADOLIBERIN I.  
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.  
FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
ACTIVITY.

FT MOD\_RES 24 24 PYROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
FT CONFLICT 16 16 W -> S (IN REF. 3).  
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 63.9%; Score 69; DB 1; Length 92;  
Best Local Similarity 70.6%; Pred. No. 0.0054;  
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CPPPSEHMSYGLRPG 17  
DB 17 CVEGSSQHSYGLRPG 33  
11:|||||

RESULT 2  
GONL\_MACMU STANDARD; PRT; 67 AA.  
AC P55247;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]  
DE (FRAGMENT).  
GN GNRH1 OR GNRH OR LHRH.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheidae; Macaca.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-HYPOTHALAMUS;  
RX MEDLINE-95124501; Pubmed-7545971;  
RA Ma Y.J., Costa M.E., Ojeda S.R.;  
RT "developmental expression of the genes encoding transforming growth  
RT factor alpha and its receptor in the hypothalamus of female rhesus  
RT macaques.";  
RL Neuroendocrinology 60:346-359(1994).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL: S75918; AAB33096.1; -  
DR INTERPRO: IPR002012; -  
DR PFAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Signal.

FT SIGNAL 1 1  
FT NON\_TER 1 5 BY SIMILARITY.  
FT CHAIN 6 67 PROGNADOLIBERIN I.  
FT PEPTIDE 6 15 GONADOLIBERIN I.  
FT PEPTIDE 19 67 GNRH-ASSOCIATED PEPTIDE I.  
FT ACT\_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
ACTIVITY (BY SIMILARITY).  
FT MOD\_RES 6 6 PYROLIDONE CARBOXYLIC ACID (BY  
SIMILARITY).  
FT MOD\_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY  
SIMILARITY).  
FT NON\_TER 67 67  
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 63.0%; Score 68; DB 1; Length 67;  
Best Local Similarity 91.7%; Pred. No. 0.0054;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 SSEHMSYGLRPG 17  
DB 4 SSQHSYGLRPG 15  
11:|||||

RESULT 3  
GONL\_MOUSE STANDARD; PRT; 90 AA.  
AC P13562;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR  
DE I].  
GN GNRH1 OR GNRH.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-87069928; Pubmed-3024317;  
RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,  
RA Phillips H.S., Nikolic K., Seeburg P.H.;  
RT "A deletion truncating the gonadotropin-releasing hormone gene is  
RT responsible for hypogonadism in the hpg mouse.";  
RL Science 234:1366-1371(1986).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL: M14872; AAA37717.1; -  
DR MGD: MGI:95789; GNRH.  
DR INTERPRO: IPR002012; -  
DR PFAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.



KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 Placenta; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 90  
 FT PEPTIDE 22 31  
 FT ACT\_SITE 24 24  
 FT MOD\_RES 22 22  
 FT MOD\_RES 31 31  
 SQ SEQUENCE 90 AA: 10337 MW: 100766FA826E4D9 CRC64;  
 Query Match 63.0%; Score 68; DB 1; Length 90;  
 Best Local Similarity 91.7%; Pred. No. 0.0072; Mismatches 0; Indels 0; Gaps 0;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 6 SSEHWSYGLRPG 17  
 DB 20 SSOHWSYGLRPG 31  
 RESULT 4  
 GONL\_PIG STANDARD; PRT; 91 AA.  
 AC P49921:  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PROCONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
 DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].  
 GN GNRH1 OR GNRH.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-HYPOTHALAMUS;  
 RA Weesner G.D., Matteri R.L., Becker B.A.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 24-33.  
 RX MEDLINE-72114303; PubMed-4946067;  
 RA Baba Y., Matsuo H., Schally A.V.;  
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.  
 RT Confirmation of the proposed structure by conventional sequential  
 RT analyses.";  
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).  
 RN [3]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE-72065376; PubMed-4942726;  
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;  
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-  
 RT phase method.";  
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).  
 RN [4]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE-72117544; PubMed-4946275;  
 RA Baba Y., Arimura A., Schally A.V.;  
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";  
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L32864; AAA1066.1; -  
 DR PIR: A01411; RHPG.  
 DR INTERPRO: IP002012; -  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 Placenta; Signal.  
 FT SIGNAL 1 23  
 FT CHAIN 24 91  
 FT PEPTIDE 24 33  
 FT ACT\_SITE 26 26  
 FT MOD\_RES 24 24  
 FT MOD\_RES 33 33  
 SQ SEQUENCE 91 AA: 10090 MW: 8340474F32DDAA99 CRC64;  
 Query Match 63.0%; Score 68; DB 1; Length 91;  
 Best Local Similarity 91.7%; Pred. No. 0.0073; Mismatches 0; Indels 0; Gaps 0;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 6 SSEHWSYGLRPG 17  
 DB 22 SSOHWSYGLRPG 33  
 RESULT 5  
 GONL\_RAT STANDARD; PRT; 92 AA.  
 AC P07490;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PROCONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
 DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR  
 DE I].  
 GN GNRH1 OR GNRH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-86094338; PubMed-2867548;  
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;  
 RT "Isolation of the gene and hypothalamic cDNA that is common precursor  
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
 RT factor in human and rat.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-89384651; PubMed-2476669;  
 RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;  
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and  
 RT hypothalamic expression.";  
 RL Mol. Endocrinol. 3:1257-1262(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-RHYMUS;  
 RX MEDLINE-93105480; PubMed-1468115;  
 RA Maler C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;  
 RT "Thymocytes express a mRNA that is identical to hypothalamic  
 RT luteinizing hormone-releasing hormone mRNA.";  
 RL Cell. Mol. Neurobiol. 12:447-454(1992).  
 RN [4]  
 RP SEQUENCE OF 1-47 FROM N.A.  
 RC TISSUE=HEART;  
 RX MEDLINE-87149087; PubMed-3547652;  
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;  
 RT "Two mammalian genes transcribed from opposite strands of the same

```

RT DNA locus.
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -.
DR EMBL; M12579; AAA41263.1; -.
DR EMBL; M31670; AAA41264.1; -.
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -.
DR EMBL; M15528; -. NOT_ANNOTATED_CDS.
DR PIR; B26173; RHRTG.
DR PIR; A48410; A48410.
DR INTERPRO; IPR002012; -.
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KM Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 1 23
FT PEPTIDE 24 33
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
SQ
Query Match 63.0%; Score 68; DB 1; Length 92;
Best Local Similarity 91.7%; Pred. No. 0.0074;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 6 SEHWSYGLRPG 17
11:|||||
DB 22 SSOHWSYGLRPG 33
RESULT 6
GONI_XENLA STANDARD; PRT; 89 AA.
ID GONI_XENLA
AC P45656;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-FORERAIN;
RX MEDLINE-94185563; PubMed-8137750;
RA Hayes W.P., Wasy S., Batley J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.

```

```

CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28040; AAA49728.1; -.
DR INTERPRO; IPR002012; -.
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KM Signal.
FT SIGNAL 1 23
FT CHAIN 1 23
FT PEPTIDE 24 33
FT PEPTIDE 24 33
FT CHAIN 37 89
FT PEPTIDE 37 85
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 89 AA; 10246 MW; 6F4F36FBAED04284 CRC64;
SQ
Query Match 56.5%; Score 61; DB 1; Length 89;
Best Local Similarity 81.8%; Pred. No. 0.061;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 7 SEHWSYGLRPG 17
::|||||
DB 23 AQHWSYGLRPG 33
RESULT 7
GONI_TUPGB STANDARD; PRT; 92 AA.
ID GONI_TUPGB
AC G95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LH-RH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRH1 OR GNRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupai.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HYPOTHALAMUS;
RX MEDLINE-97079639; PubMed-8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal."
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: U63326; AAB16837.1; -
DR INTERPRO: IPR002012; -
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 92 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 26 26 ACTIVITY.
FT MOD_RES 24 24 PYROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SEQUENCE 92 AA; 10197 MW; 4FDF2C58CF5F63B CRC64;
SQ
Query Match 56.5%; Score 61; DB 1; Length 92;
Best Local Similarity 81.8%; Pred. No. 0.063;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 7 SEHMSYGLRPG 17
Db 23 AQHMSYGLRPG 33

RESULT 8
GONL_SHEEP STANDARD; PRT; 61 AA.
ID GONL_SHEEP
AC 028588;
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
DE (FRAGMENT)
GN GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN-WESTERN RANGE; TISSUE-HYPOTHALAMUS;
RA Rodriguez R.E.; Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-10.
RX MEDLINE:72094314; PubMed=4550508;
RA Burrows R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vale W., Guillemin R.;
RT Primary structure of the ovine hypothalamic luteinizing hormone-
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-Edman degradation)";
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 002517; AAA03433.1; -
CC PIR: A93780; RSHSG.

```

```

DR INTERPRO: IPR002012; -
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 1 >61 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 3 3 ACTIVITY.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
FT SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;
SQ
Query Match 55.6%; Score 60; DB 1; Length 61;
Best Local Similarity 90.0%; Pred. No. 0.057;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 8 EHMYSYGLRPG 17
Db 1 QHMSYGLRPG 10

RESULT 9
GONL_MESAU STANDARD; PRT; 63 AA.
ID GONL_MESAU
AC 009163;
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
DE (FRAGMENT)
GN GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
RN [1]
RP Jansen H.T., Stevens P.J., Zeltner P., Lehman M.N.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
RL THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
RL HORMONES.
RN [2]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeltner P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U91938; AAB51302.1; -
CC DR INTERPRO: IPR002012; -
CC DR PFAM: PF00446; GNRH; 1.
CC PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 1 >63 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY
FT ACT_SITE 3 3 SIMILARITY).
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT MOD_RES 1 1 ACTIVITY (BY SIMILARITY).
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID (BY

```

FT MOD\_RES 10 10 SIMILARITY.  
FT AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY  
FT NON\_PER 63 63 SIMILARITY).  
SQ SEQUENCE 63 AA: 7370 MW: FC94995676F7180 CRC64;

Query Match 55.6%; Score 60; DB 1; Length 63;  
Best Local Similarity 90.0%; Pred. No. 0.058;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 8 EHWSTGLRPG 17  
:|||||  
Db 1 QHWSYGLRPG 10

RESULT 10  
GONL\_CAVPO STANDARD; PRT; 92 AA.

AC 054713;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].  
GN GNRH1 OR GNRH OR LHRH.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HARTLEY WHITE; TISSUE-HYPOTHALAMUS;  
RX MEDLINE=97462693; PubMed=9322920;  
RA Jimenez-Linan M., Rubin B.S., King J.C.;  
RT Examination of guinea pig luteinizing hormone-releasing hormone gene  
RT reveals a unique decapeptide and existence of two transcripts in the  
RL brain.";  
RL Endocrinology 138:4123-4130(1997).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AF033346; AAB87688.1; -  
DR INTERPRO: IPR002012; -  
DR PROSITE: PS00473; GNRH. 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Placenta; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 92 GONADOLIBERIN I.  
FT PEPTIDE 24 33 GNRH-ASSOCIATED PEPTIDE I.  
FT PEPTIDE 37 92 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
FT ACT\_SITE 26 26 ACTIVITY (BY SIMILARITY).  
FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY  
FT MOD\_RES 33 33 SIMILARITY).  
FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY  
FT SEQUENCE 92 AA: 10279 MW: ACF74613F456D663 CRC64;

Query Match 52.8%; Score 57; DB 1; Length 92;  
Best Local Similarity 52.9%; Pred. No. 0.21;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 CPPPSEHWSYGLRPG 17  
:|||||  
Db 17 CVENGSGYWSYGLRPG 33

RESULT 11  
GONL\_CHICK STANDARD; PRT; 92 AA.

AC P37042; P20407;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I),  
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-WHITE LECHORN;  
RX MEDLINE=94059355; PubMed=7902095;  
RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;  
RT Characterization of the chicken preprogonadotrophin-releasing  
RT hormone-I gene.";  
RL J. Mol. Endocrinol. 11:19-29(1993).  
RN [2]  
RP SEQUENCE OF 24-33.  
RC TISSUE-HYPOTHALAMUS;  
RX MEDLINE=82265778; PubMed=7050119;  
RA King J.A., Millar R.P.;  
RT "Structure of chicken hypothalamic luteinizing hormone-releasing  
RT hormone. II. Isolation and characterization.";  
RL J. Biol. Chem. 257:10729-10732(1982).  
RN [3]  
RP SEQUENCE OF 24-33.  
RC TISSUE-HYPOTHALAMUS;  
RA King J.A., Millar R.P.;  
RT "Structure of avian hypothalamic gonadotrophin-releasing hormone.";  
RL S. Afr. J. Sci. 78:124-125(1982).  
RN [4]  
RP SYNTHESIS OF 24-33.  
RX MEDLINE=82265777; PubMed=7050118;  
RA King J.A., Millar R.P.;  
RT "Structure of chicken hypothalamic luteinizing hormone-releasing  
RT hormone. I. Structural determination on partially purified  
RT material.";  
RL J. Biol. Chem. 257:10722-10728(1982).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X69491; CAA49246.1; -  
DR PIR: S33507; S33507.  
DR INTERPRO: IPR002012; -  
DR PFAM: PF00446; GNRH. 1.  
DR PROSITE: PS00473; GNRH. 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Signal.  
FT SIGNAL 1 23  
FT CHAIN 24 92  
FT PEPTIDE 24 33  
FT PEPTIDE 37 92  
FT MOD\_RES 24 24  
FT MOD\_RES 24 24

Query Match 52.8%; Score 57; DB 1; Length 92;  
Best Local Similarity 52.9%; Pred. No. 0.21;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 92 AA: 10206 MW: 61AE87EBAF508BEA CRC64;

Query Match  
 Best Local Similarity 52.8%; Score 57; DB 1; Length 92;  
 Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 Db 23 AOHWSYGLQPG 33

RESULT 12  
 GONL\_ALIMT STANDARD; PRT; 10 AA.

AC P37041; P20407;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE GONADOLIBERIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I)  
 DE (LULIBERIN I).  
 OS Alligator mississippiensis (American alligator).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Crocodylidae; Alligatorinae; Alligator.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-BRAIN;  
 RX MEDLINE-9152338; PubMed-1802082;  
 RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,  
 RA Lane V., Swanson P., Rivier J.E., Sherwood N.M.;  
 RT Primary structure of two forms of gonadotropin-releasing hormone  
 RT from brains of the American alligator (Alligator mississippiensis).  
 RL Regul. Pept. 33:105-116(1991).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC PIR: A60066; RHAQ.  
 DR INTERPRO: IPR002012;  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KW Hormone; Amidation; Hypothalamus.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA: 1172 MW: 284B23D7286B45A3 CRC64;

Query Match  
 Best Local Similarity 51.9%; Score 56; DB 1; Length 10;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 8 EHMWSYGLRPG 17  
 Db 1 OHWSYGLQPG 10

RESULT 13  
 GONL\_HAPBU STANDARD; PRT; 94 AA.

AC P51918; O93387;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)  
 DE (LH-RH I) (LULIBERIN I).  
 GN GNRH.  
 OS Haplochromis burtoni.  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;  
 CC Cichlidae; Astacotilapia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-95396797; PubMed-7667296;

RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;  
 RT "Three gonadotropin-releasing hormone genes in one organism suggest  
 RT novel roles for an ancient peptide."  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99061842; PubMed-9843638;  
 RA White R.B., Fernald R.D.;  
 RT Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression  
 RT reveals a distinct origin for GNRH-containing neurons in the  
 RT midbrain."  
 RL Gen. Comp. Endocrinol. 112:322-329(1998).  
 RN [3]  
 RP SEQUENCE OF 23-32.  
 RC TISSUE-PITUITARY;  
 RX MEDLINE-95372591; PubMed-7644702;  
 RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,  
 RA White S.A., Francis R.C., Fernald R.D., Licht P., Wardy C.,  
 RA Sherwood N.M.;  
 RT Primary structure of solitary form of gonadotropin-releasing hormone  
 RT (GNRH) in cichlid pituitary; three forms of GNRH in brain of cichlid  
 RT and pumpkinseed fish."  
 RL Regul. Pept. 57:43-53(1995).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE  
 CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALMIC-PITUITARY-  
 CC GONADAL AXIS.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS  
 CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOTHALMIC AXONS.  
 CC -1- MASS SPECTROMETRY: MW-1113.9; METHOD-MALDI; RANGE-23-32.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U31865; AAC59691.1;  
 DR EMBL: AF076961; AAC27716.1;  
 DR INTERPRO: IPR002012;  
 DR PFAM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Multigene family.  
 FT CHAIN 1 22 PROGONADOLIBERIN I.  
 FT SIGNAL 1 22 GONADOLIBERIN I.  
 FT PEPTIDE 23 32 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).  
 FT PEPTIDE 36 94 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 23 23 AMIDATION (G-33 PROVIDE AMIDE GROUP).  
 FT MOD\_RES 32 32 ENGHRTFKK -> KMDGTGRNERFL (IN REF. 1).  
 FT CONFLICT 86 94  
 SQ SEQUENCE 94 AA: 10382 MW: E57DBA8333278D7 CRC64;

Query Match  
 Best Local Similarity 51.9%; Score 56; DB 1; Length 94;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 PPSSEHWSYGLRPG 17  
 Db 18 PQGCCOHWSYGLSPG 32

RESULT 14  
 GONL\_CLAGA STANDARD; PRT; 80 AA.

AC P34339;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)

```

DE (LH-RH) (LULIBERIN I).
OS Clarias gariepinus (Sharptooth catfish) (African catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Siluriformes; Clariidae; Clarias.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE-94291651; PubMed-8020492;
RA Bogerd J., Zandbergen T., Andersson E., Goos H.;
RT "Isolation, characterization and expression of cDNAs encoding the
RT catfish-type and chicken-II-type gonadotropin-releasing-hormone
RT precursors in the African catfish."
RL Eur. J. Biochem. 222:541-549(1994).
RN [2]
RP SEQUENCE OF 25-31.
RC TISSUE-BRAIN;
RX MEDLINE-92392313; PubMed-1520292;
RA Bogerd J., Li K.W., Janssen-Dommerholt C., Goos H.;
RT "Two gonadotropin-releasing hormones from African catfish (Clarias
RT gariepinus)."
RL Biochem. Biophys. Res. Commun. 187:127-134(1992).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X78048; GAA54970.1; -.
DR EMBL: X78049; GAA54971.1; -.
DR PIR: JC1242; RHIDIS.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 21
FT CHAIN 22 80 PROGONADOLIBERIN I.
FT PEPTIDE 22 31 GONADOLIBERIN I.
FT PEPTIDE 35 80 GNRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
FT VARIANT 47 47 S -> R.
FT VARIANT 60 60 G -> R.
SQ SEQUENCE 80 AA; 8893 MW; 0BE5EE0F4FF861A CRC64;

```

Query Match 50.9%; Score 55; DB 1; Length 80;

Best Local Similarity 66.7%; Pred. No. 0.34; Mismatches 1; Indels 0; Gaps 0;

QY 6 SEHWSYGLRPG 17  
 1:|||||11  
 DB 20 SAQWHSGLNPG 31

```

RESULT 15
GONL PAGMA
ID GONL PAGMA STANDARD; PRT; 95 AA.
AC P70074;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH I) (LULIBERIN I).
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```

```

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Chrysophrys.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Okuzawa K., Graneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -----
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D86582; BAA13129.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 95 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP)
FT (BY SIMILARITY).
SQ SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;

```

Query Match 50.0%; Score 54; DB 1; Length 95;

Best Local Similarity 80.0%; Pred. No. 0.56; Mismatches 1; Indels 0; Gaps 0;

QY 8 EHWWSYGLRPG 17  
 1:|||||11  
 DB 24 QHWSYGLSPG 33

Search completed: March 2, 2001, 11:01:19  
 Job time: 434 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:55:38 ; Search time 56.87 Seconds  
(without alignments)  
20.297 Million cell updates/sec

Title: US-09-306-689-3

Perfect score: 108  
Sequence: 1 CPPPSEHWSYGLRPG 17

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	63.9	92	1	RHHUG
2	68	63.0	67	2	178541
3	68	63.0	90	1	RHMSG
4	68	63.0	92	1	RHRTG
5	61	56.5	89	2	151423
6	60	55.6	10	1	RHPGG
7	60	55.6	10	1	RHSHG
8	57	52.8	92	2	150644
9	56	51.9	10	1	RHAQ1
10	55	50.9	98	1	150739
11	55	50.9	80	1	RHTD5
12	52	48.1	2357	2	AS9249
13	51	47.2	74	2	151092
14	51	47.2	82	2	151180
15	51	47.2	82	2	151355
16	51	47.2	82	2	151365
17	51	47.2	82	2	151331
18	51	47.2	90	2	A23735
19	51	47.2	90	2	151095
20	50	46.3	497	2	T15812
21	49.5	44.8	437	2	C49274
22	47.5	44.0	305	2	T02851
23	47	43.5	10	2	A21114
24	47	43.5	190	2	T29068
25	47	43.5	226	2	T14808
26	47	43.5	322	2	S25299
27	47	43.5	325	2	T01830
28	47	43.5	327	2	T33884
29	47	43.5	361	2	S19552

30	47	43.5	396	2	A36339
31	47	43.5	679	2	A42073
32	47	43.5	769	2	156546
33	46	42.6	86	1	RHID25
34	46	42.6	164	2	C82629
35	46	42.6	252	2	T45737
36	46	42.6	404	2	T32725
37	46	42.6	412	2	B44418
38	46	42.6	543	2	H70891
39	46	42.6	551	2	A60047
40	46	42.6	756	1	B28821
41	45	41.7	122	2	A05175
42	45	41.7	164	2	G71427
43	45	41.7	169	2	T34520
44	45	41.7	282	2	C72282
45	45	41.7	335	2	D83142

## ALIGNMENTS

RESULT 1

RHHUG  
gonadoliberin precursor - human  
N:Alternate names: gonadotropin releasing hormone (GNRH); luteinizing hormone releas  
N:Contents: gonadoliberin-associated protein (GAR); progonadoliberin  
C:Species: Homo sapiens (man)  
C:Date: 17-Mar-1987 #sequence\_revision 21-Jul-1995 #text\_change 18-Jun-1999  
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718  
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.  
Nucleic Acids Res. 17, 6403-6404, 1989  
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone  
A:Reference number: S05308; MUID: 8936682  
A:Accession: S05308  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-92 <NAV>  
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gona  
A:Reference number: A94090; MUID: 86094338  
A:Accession: A26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749  
A:Experimental source: hypothalamus  
R:Seeburg, P.H.; Adelman, J.P.  
Nature 311, 666-668, 1984  
A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasin  
A:Reference number: A93342; MUID: 85012739  
A:Accession: A93342  
A:Molecule type: mRNA  
A:Residues: 1-15, 'S', 17-92 <SEE>  
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357  
A:Experimental source: placenta  
R:Tan, L.; Kousseau, P.  
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982  
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized  
A:Reference number: A90108; MUID: 83126573  
A:Accession: A90108  
A:Molecule type: protein  
A:Residues: 24-33 <TAN>  
A:Experimental source: placental trophoblasts  
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amst  
FEBS Lett. 346, 203-206, 1994  
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by  
A:Reference number: S45718; MUID: 9428397  
A:Contents: annotation; degradation pathway of synthetic hormone  
C:Genetics:  
A:Gene: GDB:GNRH; LHRH; GRH  
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760  
A:Map position: 8p21-8p11.2

FL0 protein - gard  
potassium channel  
Shaw type potassiu  
gonadoliberin II p  
hypothetical prote  
hypothetical prote  
hypothetical prote  
surface antigen -  
probable acid--COA  
adenylyl cyclase-a  
1-phosphatidylinos  
hypothetical prote  
hypothetical prote  
hypothetical prote  
oligopeptide ABC t  
hypothetical prote

A: Introns: 47/3; 79/3  
 C: Function:  
 A: Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin  
 A: Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
 C: Superfamily: gonadoliberin  
 C: Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-92/Product: progadoliberin #status predicted <PGN>  
 F:24-33/Product: gonadoliberin #status experimental <MAT>  
 F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>  
 F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen  
 F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 63.9%; Score 69; DB 1; Length 92;  
 Best Local Similarity 70.6%; Pred. No. 0.01;  
 Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CPEPSEHMSYGLRPG 17  
 11:|||||  
 Db 17 CEEGSSQHMSYGLRPG 33

RESULT 2  
 178541  
 gonadoliberin precursor - rhesus macaque (fragment)  
 N: Alternate names: luteinizing hormone releasing hormone  
 C: Species: Macaca mulatta (rhesus macaque)  
 C: Date: 02-Aug-1996 #sequence; revision 02-Aug-1996 #text-change 16-Jul-1999  
 C: Accession: I78541  
 R: Ma, Y.J.; Costa, M.E.; Ojeda, S.R.  
 Neuroendocrinology 60, 346-359, 1994  
 A: Title: Developmental expression of the genes encoding transforming growth factor alpha  
 A: Reference number: I58134; MUID:95124501  
 A: Accession: I78541  
 A: Status: preliminary; translated from GB/EMBL/DBJ  
 A: Molecule type: mRNA  
 A: Residues: 1-67 <RES>  
 A: Cross-references: GB:S75918; NID:g912831; PIDN:AAB33096.1; PID:g912832  
 C: Superfamily: gonadoliberin

Query Match 63.0%; Score 68; DB 2; Length 67;  
 Best Local Similarity 91.7%; Pred. No. 0.01;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 SSEHMSYGLRPG 17  
 11:|||||  
 Db 4 SSOHMSYGLRPG 15

RESULT 3  
 RHMSG  
 gonadoliberin precursor - mouse  
 N: Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing  
 N: Contains: gonadoliberin; gonadoliberin-associated protein (GAP)  
 C: Species: Mus musculus (house mouse)  
 C: Date: 31-Dec-1993 #sequence; revision 18-Mar-1997 #text-change 18-Jun-1999  
 C: Accession: A47578  
 R: Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolics,  
 Science 224, 1366-1371, 1986  
 A: Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo  
 A: Reference number: A47578; MUID:87069928  
 A: Accession: A47578  
 A: Molecule type: DNA  
 A: Residues: 1-90 <MAS>  
 A: Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175  
 C: Genetics:  
 A: Introns: 45/3; 77/3  
 C: Function:  
 A: Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin  
 A: Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
 C: Superfamily: gonadoliberin

C: Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-33/Product: gonadoliberin #status predicted <GAP>  
 F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>  
 F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic  
 F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 63.0%; Score 68; DB 1; Length 90;  
 Best Local Similarity 91.7%; Pred. No. 0.014;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 SSEHMSYGLRPG 17  
 11:|||||  
 Db 20 SSOHMSYGLRPG 31

RESULT 4  
 RHRTG  
 gonadoliberin precursor - rat  
 N: Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hor  
 N: Contains: gonadoliberin; prolactin release-inhibiting factor  
 C: Species: Rattus norvegicus (Norway rat)  
 C: Date: 31-Mar-1988 #sequence; revision 31-Mar-1988 #text-change 18-Jun-1999  
 C: Accession: A40147; B26173; A48410  
 R: Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.  
 Mol. Endocrinol. 3, 1257-1262, 1989  
 A: Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic  
 A: Reference number: A40147; MUID:89384661  
 A: Accession: A40147  
 A: Molecule type: DNA  
 A: Residues: 1-92 <BON>  
 A: Cross-references: GB:M31670; NID:g204447; PIDN:AAA1264.1; PID:g204448  
 R: Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
 A: Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gona  
 A: Reference number: A94090; MUID:86094338  
 A: Accession: B26173  
 A: Molecule type: mRNA  
 A: Residues: 1-92 <ADE>  
 A: Cross-references: GB:M12579; NID:g204445; PIDN:AAA1263.1; PID:g204446  
 R: Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.  
 Cell. Mol. Neurobiol. 12, 447-454, 1992  
 A: Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing horm  
 A: Reference number: A48410; MUID:93105480  
 A: Accession: A48410  
 A: Status: preliminary  
 A: Molecule type: mRNA  
 A: Residues: 1-92 <MAI>  
 A: Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060  
 A: Experimental source: thymus  
 A: Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)  
 C: Genetics:  
 A: Introns: 47/3; 79/3  
 C: Function:  
 A: Description: stimulates pituitary secretion of luteotropin and follitropin  
 A: Note: gonadoliberin-associated protein may have prolactin release inhibiting activi  
 C: Superfamily: gonadoliberin  
 C: Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-92/Product: progadoliberin #status predicted <PGN>  
 F:24-33/Product: gonadoliberin #status predicted <GAP>  
 F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>  
 F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic  
 F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 63.0%; Score 68; DB 1; Length 92;  
 Best Local Similarity 91.7%; Pred. No. 0.014;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 SSEHMSYGLRPG 17  
 11:|||||



Db 22 SSOHWSYGLRPG 33

RESULT 5  
151423  
gonadoliberin precursor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C:Accession: 151423  
R:Hayes, W.P.; Wray, S.; Battey, J.F.  
Endocrinology 134, 1835-1845, 1994  
A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain  
A:Reference number: 151423; MUID:94185563  
A:Accession: 151423  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-89 <HAY>  
A:Cross-references: GB:L28040; NID:g496291; PIDN:AAA9728.1; PID:g496292  
C:Genetics:  
A:Gene: GnRH-I  
C:Superfamily: gonadoliberin

Query Match 56.5%; Score 61; DB 2; Length 89;  
Best Local Similarity 81.8%; Pred. No. 0.12;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Oy 7 SEHWSYGLRPG 17  
:::|||||  
Db 23 AQHWSYGLRPG 33

RESULT 6  
RHPG  
gonadoliberin - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 13-Jul-1991 #sequence\_revision 13-Jul-1991 #text\_change 18-Mar-1997  
C:Accession: A01411  
R:Baba, Y.; Matsuo, H.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 44, 459-463, 1971  
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the  
A:Reference number: A90172; MUID:72114303  
A:Accession: A01411  
A:Molecule type: protein  
A:Residues: 1-10 <BAB>  
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 45, 822-827, 1971  
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase method  
A:Reference number: A90176; MUID:7206376  
A:Accession: A01411  
A:Contents: annotation  
A:Note: the synthetic and natural hormones have the same physicochemical and biological  
R:Baba, Y.; Arimura, A.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 45, 483-487, 1971  
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.  
A:Reference number: A90175; MUID:72117544  
A:Contents: annotation  
A:Note: Trp-3 appears to be essential for biological activity  
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle-stimulating hormone  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 55.6%; Score 60; DB 1; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.016;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Oy 8 EHWWSYGLRPG 17  
:|||||  
Db 1 QHWSYGLRPG 10

RESULT 7

RHSHG  
gonadoliberin - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Mar-1997  
C:Accession: A93780; A01411  
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;  
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972  
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor  
A:Reference number: A93780; MUID:72094314  
A:Accession: A93780  
A:Molecule type: protein  
A:Residues: 1-10 <SHR>  
A:Note: The natural and synthetic hormones have the same biological activity  
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle-stimulating hormone  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 55.6%; Score 60; DB 1; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.016;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Oy 8 EHWWSYGLRPG 17  
:|||||  
Db 1 QHWSYGLRPG 10

RESULT 8  
150644  
gonadoliberin I precursor - chicken  
N:Alternate names: gonadotropin-releasing hormone I  
C:Species: Gallus gallus (chicken)  
C>Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999  
C:Accession: I50644; S33507  
R:Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.  
J. Mol. Endocrinol. 11, 19-29, 1993  
A:Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene  
A:Reference number: I50644; MUID:94059355  
A:Accession: I50644  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-92 <DU2>  
A:Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA49246.1; PID:g311612  
C:Genetics:  
A:Insertions: 47/3; 79/3  
C:Superfamily: gonadoliberin

Query Match 52.8%; Score 57; DB 2; Length 92;  
Best Local Similarity 72.7%; Pred. No. 0.41;  
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Oy 7 SEHWSYGLRPG 17  
:::|||||  
Db 23 AQHWSYGLRPG 33

RESULT 9  
RHA01  
gonadoliberin I - American alligator  
N:Alternate names: gonadotropin-releasing hormone I  
C:Species: Alligator mississippiensis (American alligator)  
C>Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Mar-1997  
C:Accession: A60066  
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan  
Regul. Pept. 33, 105-116, 1991  
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains  
A:Reference number: A60066; MUID:91352338  
A:Accession: A60066

A:Molecule type: protein  
 A:Residues: 1-10 <LOV>  
 C:Superfamily: gonadoliberin  
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
 F:1-21/Domain: signal sequence #status predicted <Sig>  
 F:22-31/Product: gonadoliberin I #status experimental <MAT1>  
 F:35-80/Product: gonadoliberin I-associated protein #status predicted <MAT2>  
 F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi  
 F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 51.9%; Score 56; DB 1; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 0.054;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 8 EHMSYGLRPG 17  
 :|||||:|  
 Db 1 OHMSYGLPG 10

RESULT 10  
 150739

gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)  
 C:Species: Haplochromis burtoni

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: 150739

R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.  
 Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995

A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles  
 A:Reference number: 150739; MUID:95396797

A:Accession: 150739

A:Status: preliminary; translated from GB/EMBL/DBU

A:Molecule type: mRNA

A:Residues: 1-98 <WH>

A:Cross-references: EMBL:U01865; NID:9905398; PIDN:MAC59691.1; PID:9905399

C:Superfamily: gonadoliberin

Query Match 51.9%; Score 56; DB 2; Length 98;  
 Best Local Similarity 60.0%; Pred. No. 0.6;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 PPSEHMSYGLRPG 17  
 :|||||:|  
 Db 18 POGCCOHMSTGLSPG 32

RESULT 11  
 RHID15

gonadoliberin I precursor - sharpshooth catfish

N:Alternate names: gonadoliberin, catfish-type; gonadotropin-releasing hormone I (GnRH-I)

C:Species: Clarias gariepinus (sharpshooth catfish)

C>Date: 30-Sep-1993 #sequence\_revision 18-Mar-1997 #text\_change 18-Jun-1999

C:Accession: S45602; J01242; S42936; S42937

R:Bogerd, J.; Zandbergen, T.; Andersson, E.; Goos, H.  
 Eur. J. Biochem. 222, 541-549, 1994

A:Title: Isolation, characterization and expression of cDNAs encoding the catfish-type a

A:Reference number: S45600; MUID:94291651

A:Molecule type: mRNA

A:Residues: 1-80 <BOG1>

A:Cross-references: EMBL:X78049; NID:9459433; PIDN:CAAS4971.1; PID:9459434

A:Molecule type: protein

A:Accession: S45601

A:Molecule type: mRNA

A:Residues: 1-46, 'S', 48-59, 'G', 61-80 <BOG2>

A:Cross-references: EMBL:X78048; NID:9459431; PIDN:CAAS4970.1; PID:9459432

A:Note: gonadoliberin I-associated protein form II, presumed to be a polymorphic form

R:Bogerd, J.; Li, K.W.; Janssen-Dommerholt, C.; Goos, H.  
 Biochem. Biophys. Res. Commun. 187, 127-134, 1992

A:Title: Two gonadotropin-releasing hormones from African catfish (Clarias gariepinus).  
 A:Reference number: J01242; MUID:92392313

A:Accession: J01242

A:Molecule type: protein

A:Residues: 22-31 <BOG3>

A:Experimental source: brain  
 C:Superfamily: gonadoliberin  
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
 F:1-21/Domain: signal sequence #status predicted <Sig>  
 F:22-31/Product: gonadoliberin I #status experimental <MAT1>  
 F:35-80/Product: gonadoliberin I-associated protein #status predicted <MAT2>  
 F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi  
 F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 50.9%; Score 55; DB 1; Length 80;  
 Best Local Similarity 66.7%; Pred. No. 0.66;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 6 SSEHMSYGLRPG 17  
 :|||||:|  
 Db 20 SAQHMSHGLNPG 31

RESULT 12  
 A59249

Class VII unconventional myosin - slime mold (Dictyostelium discoideum)

C:Species: Dictyostelium discoideum

C>Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 08-Sep-2000

C:Accession: A59249

R:Titus, M.A.; Kuspa, A.; Loomis, W.F.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 9446-9450, 1994

A:Title: Discovery of myosin genes by physical mapping in Dictyostelium.

A:Reference number: Z20873; MUID:95023928

A:Accession: A59249

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-2357 <LT1>

A:Cross-references: GB:U35321; NID:96226760; PIDN:AAF06035.1; PID:96226761

A:Experimental source: gene myoI; product class VII unconventional myosin

R:Titus, M.A.  
 Curr. Biol. 9, 1297-1303, 1999

A:Title: A class VII unconventional myosin is required for phagocytosis.

A:Reference number: A59249

A:Contents: annotation

C:Genetics:

A:Gene: myoI

C:Superfamily: myosin motor domain homology

F:16-676/Domain: myosin motor domain homology <MMO>

Query Match 48.1%; Score 52; DB 2; Length 2357;  
 Best Local Similarity 60.0%; Pred. No. 59;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 PPPSEHMSYGLRPG 16  
 :|||||:|  
 Db 1027 PPPSSSSMDFGLRP 1041

RESULT 13  
 151092

gonadotropin releasing hormone - chinook salmon (fragment)

C:Species: Oncorhynchus tshawytscha (chinook salmon)

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Aug-1997

C:Accession: 151092

R:Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.  
 Mol. Cell. Endocrinol. 84, 167-174, 1992

A:Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.

A:Reference number: 151040; MUID:92267241

A:Accession: 151092

A:Status: preliminary; translated from GB/EMBL/DBU

A:Molecule type: DNA

A:Residues: 1-74 <KLH>

A:Cross-references: EMBL:X79711; NID:9499322; PID:9499323

C:Genetics:

A:Gene: GnRH

A:Introns: 38/3; 65/3

Job time: 259 sec

Query Match 47.2%; Score 51; DB 2; Length 74;  
 Best Local Similarity 72.7%; Pred. No. 2.1;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 1:||||| 11  
 DB 15 SQHWSYGLWLP 25

## RESULT 14

151180

gonadotropin-releasing hormone - cherry salmon

C:Species: Oncorhynchus masou (cherry salmon)

C&gt;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Aug-1997

C:Accession: 151180

R:Suzuki, M.; Hyodo, S.; Kobayashi, M.; Aida, K.; Urano, A.

J. Mol. Endocrinol. 9, 73-82, 1992.

A&gt;Title: Characterization and localization of mRNA encoding the salmon-type gonadotrophin

A:Reference number: 151180; MUID:92284893

A:Accession: 151180

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-82 &lt;SU2&gt;

A:Cross-references: GB:S44614; NID:9254824; PID:9254825

Query Match 47.2%; Score 51; DB 2; Length 82;  
 Best Local Similarity 72.7%; Pred. No. 2.3;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 1:||||| 11  
 DB 23 SQHWSYGLWLP 33

## RESULT 15

151355

gonadotropin releasing hormone - Atlantic salmon

C:Species: Salmo salar (Atlantic salmon)

C&gt;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: 151355

R:Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.

Mol. Cell. Endocrinol. 84, 167-174, 1992.

A&gt;Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.

A:Reference number: 151040; MUID:92267241

A:Accession: 151355

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-82 &lt;KLU&gt;

A:Cross-references: EMBL:X79709; NID:9499341; PID:9499342

A:Accession: 151355

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-82 &lt;KL2&gt;

A:Cross-references: EMBL:X74957; NID:9402786; PIDN:CA52912.1; PID:9402787

A:Gene: GnRH

A:Introns: 46/3; 73/3

Query Match 47.2%; Score 51; DB 2; Length 82;  
 Best Local Similarity 72.7%; Pred. No. 2.3;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17  
 1:||||| 11  
 DB 23 SQHWSYGLWLP 33

Search completed: March 2, 2001, 10:55:39

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:10 ; Search time 97.13 Seconds  
(without alignments)  
16.894 Million cell updates/sec

Title: US-09-306-689-2  
Perfect score: 58  
Sequence: 1 XHMSYXLRPEXXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_15:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.podent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	87.9	91	13	Q9PRH0
2	45	77.6	87	13	Q9Y126
3	45	77.6	95	13	Q73812
4	45	77.6	99	13	Q91A10
5	43	74.1	90	13	Q91AU2
6	42	72.4	565	2	Q9KML1
7	39	67.2	508	10	Q82588
8	39	67.2	828	10	Q9SZR5
9	38	65.5	33	13	Q9W7G0
10	38	65.5	33	13	Q9PT34
11	38	65.5	82	13	Q92094
12	38	65.5	82	13	Q9W7G1
13	38	65.5	82	13	Q91B00
14	38	65.5	82	13	Q91B00
15	38	65.5	88	13	Q9PSY9
16	38	65.5	90	13	Q91A09
17	38	65.5	205	2	Q9S299
18	37	63.8	954	10	Q22613
19	37	63.8	965	10	Q9M4N4

20	37	63.8	966	10	Q9M4N3	Q9M4N3 medicago tr
21	36	62.1	289	2	Q9X2P4	Q9X2P4 mycobacteri
22	36	62.1	321	10	Q9S064	Q9S064 papaver som
23	36	62.1	379	10	Q9LV12	Q9LV12 arabidopsis
24	36	62.1	425	5	Q9U9U4	Q9U9U4 drosophila
25	36	62.1	612	10	Q9M8N3	Q9M8N3 arabidopsis
26	36	62.1	954	10	Q9SMW2	Q9SMW2 nicotiana p
27	36	62.1	954	10	Q9SAJ1	Q9SAJ1 arabidopsis
28	36	62.1	956	10	Q9LV11	Q9LV11 arabidopsis
29	35	60.3	32	5	Q9U340	Q9U340 caenorhabdi
30	35	60.3	82	1	Q28696	Q28696 archaeoglob
31	35	60.3	148	2	Q51920	Q51920 streptomyce
32	35	60.3	155	3	Q9UR56	Q9UR56 kluyveromyc
33	35	60.3	155	8	Q9XMD4	Q9XMD4 kluyveromyc
34	35	60.3	161	2	Q9RVDB	Q9RVDB deinococcus
35	35	60.3	220	5	Q21256	Q21256 caenorhabdi
36	35	60.3	479	2	Q9L0N5	Q9L0N5 streptomyce
37	35	60.3	485	2	Q9L0N4	Q9L0N4 streptomyce
38	35	60.3	485	2	Q9L0N3	Q9L0N3 streptomyce
39	35	60.3	526	2	Q9J041	Q9J041 chlamydia p
40	35	60.3	567	4	Q13208	Q13208 homo sapien
41	35	60.3	691	10	Q23482	Q23482 arabidopsis
42	35	60.3	711	4	Q14870	Q14870 homo sapien
43	35	60.3	1072	6	Q77690	Q77690 bos taurus
44	35	60.3	1232	10	Q9M264	Q9M264 arabidopsis
45	34	58.6	119	1	Q9Y9R2	Q9Y9R2 aetopyrum p

#### ALIGNMENTS

```
RESULT 1
Q9PRH0 PRELIMINARY: PRT; 91 AA.
AC Q9PRH0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DI 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DI 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE PREPRO-MGNRH PRECURSOR.
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguillidae; Anguilla.
OX NCBI_TaxID=9937;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Okubo K.; Suetake H.; Aida K.;
RT "Expression of two gonadotropin-releasing hormone (GnRH) precursor
RT genes in various tissues of the Japanese eel and evolution of GnRH." ;
RL Zool. Sci. 16:471-478(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Okubo K.; Suetake H.; Aida K.;
RT hormone (prepro-mgnrh) mRNA is present in the brain and various
RT peripheral tissues of the Japanese eel." ;
RL Zool. Sci. 16:645-651(1999).
DR EMBL: AB026989; BAA82608.1; -
DR EMBL: AB026991; BAA83597.1; -
DR INTERPRO: IPR002012; -
DR PIRAM: PF00446; GNRH.1.
DR PROSITE: PS00473; GNRH.1.
KW signal.
FT CHAIN 1 22 POTENTIAL.
FT CHAIN 23 32 GNRH.
FT CHAIN 33 91 GNRH ASSOCIATED PEPTIDE.
SQ SEQUENCE 91 AA: 9893 MW: 84135.00
Query Match 87.9%; Score 51; DB 13; Length 91;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

OY 2 HWSYXLRPG 10  
 |||||  
 Db 24 HWSYGLSPG 32

## RESULT 2

ID 09Y126 PRELIMINARY; PRT; 87 AA.

AC 09Y126; PRELIMINARY; PRT; 87 AA.

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)

OS Sparus aurata (Gilthead sea bream).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;

OC Sparidae; Sparus.

NCBI\_TaxID=8175;

RP SEQUENCE FROM N.A.

RC TISSUE=OVARY;

RA Nabissi M.;

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND

CC POLICIE-STIMULATING HORMONES.

CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.

DR EMBL; AF046801; AAD02427.1; -

DR INTERPRO; IPR002012; -

DR PFAM; PF00446; GNRH; 1.

DR PROSITE; PS00473; GNRH; 1.

KM Hormone; Amidation.

FT NON\_TER 1

FT NON\_TER 87

FT NON\_TER 87

SO SEQUENCE 87 AA; 9871 MW; 0D2463533D96782A CRC64;

Query Match Best Local Similarity 77.6%; Score 45; DB 13; Length 87;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 Db 22 HWSYGLSPG 30

## RESULT 3

ID 073812 PRELIMINARY; PRT; 95 AA.

AC 073812; PRELIMINARY; PRT; 95 AA.

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).

OS Morone saxatilis (Striped bass).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;

OC Moronidae; Morone.

NCBI\_TaxID=34816;

RP SEQUENCE FROM N.A.

RA Chow M.M., Kight K.E., Gotthalf Y., Alok D., Zohar Y.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND

CC POLICIE-STIMULATING HORMONES.

CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.

DR EMBL; AF056314; AAD03817.1; -

DR INTERPRO; IPR002012; -

DR PFAM; PF00446; GNRH; 1.

DR PROSITE; PS00473; GNRH; 1.

DR PRODOM; PD005581; -; 1.  
 KM Hormone; Amidation.  
 SO SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match Best Local Similarity 77.6%; Score 45; DB 13; Length 95;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 Db 24 HWSYGLSPG 32

## RESULT 4

ID 091A10 PRELIMINARY; PRT; 99 AA.

AC 091A10; PRELIMINARY; PRT; 99 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE GONADOTROPIN-RELEASING HORMONE SEABREAM ISOFORM.

OS Dicentrarchus labrax (European sea bass).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;

OC Moronidae; Dicentrarchus.

NCBI\_TaxID=13489;

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;

RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanuy S.;

RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;

RT "Differential expression of three different prepro-GNRH

RT (gonadotropin-releasing hormone) messengers in the brain of the

RT European sea bass (Dicentrarchus labrax).";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;

RA Zmora N., Zohar Y., Elizur A.;

RT "3 GNRH form in the seabass Dicentrarchus labrax.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF224279; AAF62898.1; -

SO SEQUENCE 99 AA; 10758 MW; EC8AEEC93CC02904 CRC64;

Query Match Best Local Similarity 77.6%; Score 45; DB 13; Length 99;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 Db 28 HWSYGLSPG 36

## RESULT 5

ID 091A02 PRELIMINARY; PRT; 90 AA.

AC 091A02; PRELIMINARY; PRT; 90 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE GONADOTROPIN-RELEASING HORMONE.

OS Rana dybowskii (Frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.

NCBI\_TaxID=71582;

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;

RA Yoo M.S., Kang H.M., Choi H.S., Chun S.Y., Troskie B., Millar R.P.;

RA Kwon H.B.;

RT "Molecular Cloning, Distribution and Pharmacological Characterization  
of a Novel Gonadotropin-Releasing Hormone([trp8]GnRH) in Frog Brain.";  
RL Mol. Cell. Endocrinol. 0:0-0(2000).  
DR EMBL: AF139911; AAF44343.1; -  
SQ SEQUENCE 90 AA; 10368 MW; C3D573E78B52ABFA CRC64;

Query Match 74.1%; Score 43; DB 13; Length 90;  
Best Local Similarity 77.8%; Pred. No. 0.69;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
DB 26 HWSYGLWPG 34

RESULT 6  
O9KM11 PRELIMINARY; PRT; 565 AA.

AC 09KM11:  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE HYPOTHETICAL PROTEIN VCA0578.  
GN VCA0578.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE=204066833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
cholerae";  
RL Nature 406:477-483(2000).  
DR EMBL: AE004388; AAF96480.1; -  
DR TIGR: VCA0578; -  
SQ SEQUENCE 565 AA; 64837 MW; 36GD8479E2ADE303 CRC64;

Query Match 72.4%; Score 42; DB 2; Length 565;  
Best Local Similarity 66.7%; Pred. No. 6.5;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
DB 179 HWRFLRPG 187

RESULT 7  
O82588 PRELIMINARY; PRT; 508 AA.  
AC 082588:  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE F1104.7 PROTEIN (FRAGMENT).  
GN F1104.7.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;

RA WASHU;  
RT "The A. thaliana Genome Sequencing Project";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Abu-Threideh J., Stoneking T., Langston Y., Trevasakis E.,  
RT "The sequence of A. thaliana F1104.";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Waterston R.;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF096370; AAC62783.1; -  
DR INTERPRO: IPR002885; -  
DR PFM: PF01535; PPR: 9.  
FT NON\_TER 508 508  
SQ SEQUENCE 508 AA; 56879 MW; EB1449A49067AEDD CRC64;

Query Match 67.2%; Score 39; DB 10; Length 508;  
Best Local Similarity 75.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WSYXLRPG 10  
DB 80 WCYSLRPG 87

RESULT 8  
O9SZR5 PRELIMINARY; PRT; 828 AA.

AC 09SZR5:  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE HYPOTHETICAL 93.2 KDA PROTEIN.  
GN F27B13.190 OR ATG629950.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Bancroft I.,  
RA Mewes H.W., Mayer K.F.X., Schueller C.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Zimmermann W., Gruenisen A., Wambutt R., Kalicki J., Wohlmann P.,  
RA Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL050352; CAB43670.1; -  
DR EMBL: AL161575; CAB9753.1; -  
DR INTERPRO: IPR000195; -  
RV Hypothetical protein.  
SQ SEQUENCE 828 AA; 93177 MW; CAB050A332BC5706 CRC64;

Query Match 67.2%; Score 39; DB 10; Length 828;  
Best Local Similarity 66.7%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 DB 114 HWSYGWLP 122

RESULT 9  
 O9W7G0 PRELIMINARY; PRT; 33 AA.  
 AC O9W7G0;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN) (FRAGMENT).  
 GN GNRH2.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99312119; PubMed=10385393;  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RT "Regulation and expression of gonadotropin-releasing hormone gene  
 RT differs in brain and gonads in rainbow trout.";  
 RL Endocrinology 140:3012-3024(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
 DR EMBL: AF110993; ADD43463.1; -.  
 DR INTERPRO: IPR002012; -.  
 DR PFW: PFO0446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 DR KW Hormone; Amidation.  
 FT NON\_TER 33  
 FT SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;

Query Match 65.5%; Score 38; DB 13; Length 33;  
 Best Local Similarity 66.7%; Pred. No. 2.1;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 DB 25 HWSYGWLP 33

RESULT 10  
 O9PT34 PRELIMINARY; PRT; 33 AA.  
 AC O9PT34;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE GONADOTROPIN-RELEASING HORMONE (FRAGMENT).  
 GN GNRH1.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99312119; PubMed=10385393;  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RT "Regulation and expression of gonadotropin-releasing hormone gene  
 RT differs in brain and gonads in rainbow trout.";  
 RL Endocrinology 140:3012-3024(1999).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Von Schalburg K.R., Sherwood N.M.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC EMBL: AF110533; ADD43461.1; -.  
 DR INTERPRO: IPR002012; -.  
 DR PFW: PFO0446; GNRH; 1.  
 DR PROSITE: PS00256; AKH; UNKNOWN\_1.  
 DR KW Hormone; Amidation.  
 FT NON\_TER 33  
 FT SEQUENCE 33 AA; 3741 MW; 1FE1355E742B7EBB CRC64;

Query Match 65.5%; Score 38; DB 13; Length 33;  
 Best Local Similarity 66.7%; Pred. No. 2.1;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 DB 25 HWSYGWLP 33

RESULT 11  
 O9Z094 PRELIMINARY; PRT; 82 AA.  
 AC O9Z094;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN) (FRAGMENT).  
 GN PREPRO-GNRH-1.  
 OS Oncorhynchus nerka (Sockeye salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8023;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=NIKKO; TISSUE=BRAIN;  
 RX MEDLINE=96020547; PubMed=8546809;  
 RA Ashihara M., Suzuki M., Kubokawa K., Yoshitake Y., Kobayashi M.,  
 RA Urano A., Aida K.;  
 RT "Two differing precursor genes for the salmon-type gonadotropin-  
 RT releasing hormone exist in salmonids.";  
 RL J. Mol. Endocrinol. 15:1-9(1995).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND  
 CC FOLLICLE-STIMULATING HORMONES.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
 DR EMBL: D31868; BAA06666.1; -.  
 DR INTERPRO: IPR002012; -.  
 DR PFW: PFO0446; GNRH; 1.  
 DR PROSITE: PS00256; AKH; UNKNOWN\_1.  
 DR KW Hormone; Amidation.  
 FT SIGNAL 1 23  
 FT CHAIN 24 33  
 FT CHAIN 37 82  
 FT SEQUENCE 82 AA; 9126 MW; C64044EA521B2BBB CRC64;

Query Match 65.5%; Score 38; DB 13; Length 82;  
 Best Local Similarity 66.7%; Pred. No. 5.2;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 DB 25 HWSYGWLP 33



```

RESULT 12
O9W7G1 PRELIMINARY: PRT: 82 AA.
AC O9W7G1:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).
GN GNRH.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
RT differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
DR EMBL; AF110992; AAD3462.1; -
DR INTERPRO: IPR002012; -
DR INTERPRO: IPR002047; -
DR PFM: PF00446; GNRH: 1.
DR PROSITE: PS00256; ARH: UNKNOWN_1.
DR PROSITE: PS00473; GNRH: 1.
KW Hormone; Amidation.
SQ SEQUENCE 82 AA; 9232 MW; 7595B4FCC65FDFD6 CRC64;

Query Match 65.5%; Score 38; DB 13; Length 82;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
Db 25 HWSYXLRPG 33

RESULT 13
O91800 PRELIMINARY: PRT: 82 AA.
AC O91800:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PRO-SGNRH-1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Ferriere F., Bailhache T., Jégo P.;
RT "Oncorhynchus mykiss sgnrh-I cDNA from brain.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF232212; AAF91280.1; -
SQ SEQUENCE 82 AA; 9198 MW; 7595A0B896556A69 CRC64;

Query Match 65.5%; Score 38; DB 13; Length 82;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 2 HWSYXLRPG 10
Db 25 HWSYXLRPG 33

RESULT 14
O918P9 PRELIMINARY: PRT: 82 AA.
AC O918P9:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PRO-SGNRH-II.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Ferriere F., Bailhache T., Jégo P.;
RT "Oncorhynchus mykiss sgnrh-II cDNA in the brain.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF232213; AAF91281.1; -
SQ SEQUENCE 82 AA; 9203 MW; 8053P4F221A0F08 CRC64;

Query Match 65.5%; Score 38; DB 13; Length 82;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
Db 25 HWSYXLRPG 33

RESULT 15
O9PSY9 PRELIMINARY: PRT: 88 AA.
AC O9PSY9:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE SGNRH (FRAGMENT).
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Nabissi M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF046799; AAD02425.1; -
DR INTERPRO: IPR002012; -
DR PFM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
FT NON_TER 88
SQ SEQUENCE 88 AA; 9788 MW; F7EB868C2FBD19F CRC64;

Query Match 65.5%; Score 38; DB 13; Length 88;
Best Local Similarity 66.7%; Pred. No. 5.6;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Query Match 65.5%; Score 38; DB 13; Length 82;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Fri Mar 2 12:19:26 2001

us-09-306-689-2.rspt

Search completed: March 2, 2001, 11:00:45  
Job time: 455 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:51:20 ; Search time 56.87 Seconds  
(without alignments)  
16.715 Million cell updates/sec

Title: US-09-306-689-2  
Perfect score: 58  
Sequence: 1 XHMSYXLRPGXXXX 14

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	87.9	10	1	RHPGC gonadoliberin - pi
2	51	87.9	10	1	RHSHG gonadoliberin - sh
3	51	87.9	67	2	I78541 gonadoliberin prec
4	51	87.9	89	2	I51423 gonadoliberin prec
5	51	87.9	90	1	RHMSG gonadoliberin prec
6	51	87.9	92	1	RHHUG gonadoliberin prec
7	51	87.9	92	1	RHRTG gonadoliberin prec
8	47	81.0	10	1	RHAQ1 gonadoliberin I -
9	47	81.0	92	2	I50644 gonadoliberin I pr
10	45	77.6	98	2	I50739 gonadotropin-relea
11	42	72.4	551	2	E64728 yabn protein - esc
12	42	72.4	565	2	G82443 conserved hypotet
13	41	70.7	80	1	RHID1S gonadoliberin I pr
14	39	67.2	508	2	T01937 hypothetical prote
15	39	67.2	828	2	T08556 hypothetical prote
16	38	65.5	10	2	A21114 gonadoliberin - ch
17	38	65.5	74	2	I51092 gonadotropin relea
18	38	65.5	82	2	I51180 gonadotropin-relea
19	38	65.5	82	2	I51355 gonadotropin relea
20	38	65.5	82	2	I51365 gonadotropin relea
21	38	65.5	82	2	I51331 gonadotropin relea
22	38	65.5	90	2	A23735 gonadoliberin prec
23	38	65.5	90	2	I51095 gonadoliberin prec
24	38	65.5	205	2	T36753 hypothetical prote
25	37	63.8	10	2	A49187 gonadotropin-relea
26	37	63.8	584	2	U01229 cellulase (EC 3.2.
27	36	62.1	954	2	S60301 H+-transporting AT
28	35	60.3	82	2	G69446 hypothetical prote
29	35	60.3	161	2	C75438 hypothetical prote

30	35	60.3	220	2	T23358 hypothetical prote
31	35	60.3	390	2	S54026 ribosomal protein
32	35	60.3	477	1	C6CLAM cellulase (EC 3.2.
33	35	60.3	492	2	C83592 hypothetical prote
34	35	60.3	526	2	G72091 4-alpha-glucanot
35	35	60.3	611	2	S57527 DNA helicase II -
36	35	60.3	691	2	D71430 hypothetical prote
37	35	60.3	711	1	A47136 macrophage-stimula
38	35	60.3	1232	2	T47993 hypothetical prote
39	34	58.6	119	2	F72531 hypothetical prote
40	34	58.6	138	2	G71506 ribosomal protein
41	34	58.6	316	2	A53440 aldose reductase h
42	34	58.6	351	2	T28864 hypothetical prote
43	34	58.6	486	2	S55146 probable ribosomal
44	34	58.6	676	2	A45900 complement C3b rec
45	34	58.6	972	2	T47922 anthranilate phosph

## ALIGNMENTS

```

RESULT 1
RHPGC
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A:Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match      87.9%  Score 51:  DB 1:  Length 10:
Best Local Similarity 88.9%  Pred. No. 0.0019:
Matches      8:  Conservative      0:  Mismatches      1:  Indels      0:  Gaps      0:

OY      2  HMSYXLRPG 10
DB      2  HMSYXLRPG 10

RESULT 2
RSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Butcher, R.; Butcher, M.; Amos, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780

```

A:Molecule type: protein  
A:Residues: 1-10 <BR>  
A:Note: the natural and synthetic hormones have the same biological activity  
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.0019;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||||  
Db 2 HWSYGLRPG 10

RESULT 3  
178541  
gonadoliberin precursor - rhesus macaque (fragment)  
N:Alternate names: luteinizing hormone releasing hormone  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: I78541  
R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.  
Neuroendocrinology 60, 346-359, 1994  
A:Title: Developmental expression of the genes encoding transforming growth factor alpha  
A:Reference number: I58134; M0ID:95124501  
A:Accession: I78541  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-67 <RES>  
A:Cross-references: GB:S75918; NID:g912831; PIDN:AA33096.1; PID:g912832  
C:Superfamily: gonadoliberin

Query Match 87.9%; Score 51; DB 2; Length 67;  
Best Local Similarity 88.9%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||||  
Db 7 HWSYGLRPG 15

RESULT 4  
151423  
gonadoliberin precursor - African clawed frog  
N:Alternate names: luteinizing hormone releasing hormone  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C:Accession: I51423  
R:Hayes, W.P.; Wray, S.; Battey, J.F.  
Endocrinology 134, 1835-1845, 1994  
A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain  
A:Reference number: I51423; M0ID:94185563  
A:Accession: I51423  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-89 <HAY>  
A:Cross-references: GB:L28040; NID:g496291; PIDN:AAA9728.1; PID:g496292  
C:Genetics:  
A:Gene: GnRH-I  
C:Superfamily: gonadoliberin

Query Match 87.9%; Score 51; DB 2; Length 89;  
Best Local Similarity 88.9%; Pred. No. 0.016;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

Db 25 HWSYGLRPG 33  
|||||

RESULT 5  
RHMSG  
gonadoliberin precursor - mouse  
N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1993 #sequence\_revision 18-Mar-1997 #text\_change 18-Jun-1999  
C:Accession: A47578  
R:Nason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli  
Science 234, 1366-1371, 1986  
A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible for the hypogonadism in the mouse  
A:Reference number: A47578; M0ID:87069928  
A:Accession: A47578  
A:Molecule type: DNA  
A:Residues: 1-90 <MAS>  
A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175  
C:Genetics:  
A:Introns: 45/3; 77/3  
C:Function:  
A:Description: gonadoliberin stimulates pituitary secretion of luteinizing hormone and follicle stimulating hormone  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:22-31/Product: gonadoliberin #status predicted <GLR>  
F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:22/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predicted <RES>  
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form) #status predicted <RES>

Query Match 87.9%; Score 51; DB 1; Length 90;  
Best Local Similarity 88.9%; Pred. No. 0.016;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||||  
Db 23 HWSYGLRPG 31

RESULT 6  
RHMSG  
gonadoliberin precursor - human  
N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)  
C:Species: Homo sapiens (man)  
C:Date: 17-Mar-1987 #sequence\_revision 21-Jul-1995 #text\_change 18-Jun-1999  
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718  
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.  
Nucleic Acids Res. 17, 6403-6404, 1989  
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone gene  
A:Reference number: S05308; M0ID:89366682  
A:Accession: S05308  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-92 <HAY>  
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and growth hormone releasing hormone  
A:Reference number: A94090; M0ID:86094338  
A:Accession: A26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749  
R:Seeburg, P.H.; Adelman, J.P.  
Nature 311, 666-668, 1984  
A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing hormone  
A:Reference number: A93342; M0ID:85012739

A:Accession: A93342  
A:Molecule type: mRNA  
A:Residues: 1-15, 'S', 17-92 <SEB>  
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357  
A:Experimental source: placenta  
R:Tan, L.; Rousseau, P.  
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982  
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in  
A:Reference number: A90108; MUID:83126573  
A:Accession: A90108  
A:Molecule type: protein  
A:Residues: 24-33 <TAN>  
A:Experimental source: placental trophoblasts  
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterde  
FEBS Lett. 346, 203-206, 1994  
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th  
A:Reference number: S45718; MUID:94283597  
A:Contents: annotation; degradation pathway of synthetic hormone  
C:Genetics:  
A:Gene: GDB:GNRH; LHRH; GRH  
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760  
A:Map position: bp21-8p11.2  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin  
C:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
F:1-23/Dominant: signal sequence #status predicted <SIG>  
F:24-92/Product: progonaoliberin #status predicted <PCN>  
F:24-33/Product: gonadoliberin #status predicted <MAT>  
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 87.9%; Score 51; DB 1; Length 92;  
Best Local Similarity 88.9%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 HWSYGLRPG 10  
Db 25 HWSYGLRPG 33

RESULT 7  
RHRGS  
gonadoliberin precursor - rat  
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormo  
N:Contents: gonadoliberin; prolactin release-inhibiting factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1999  
C:Accession: A40147; B26173; A48410  
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.  
Mol. Endocrinol. 3, 1257-1262, 1989  
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex  
A:Reference number: A40147; MUID:89384661  
A:Accession: A40147  
A:Molecule type: DNA  
A:Residues: 1-92 <BON>  
A:Cross-references: GB:M1670; NID:g204447; PIDN:AAA41264.1; PID:g204448  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot  
A:Reference number: A94090; MUID:86094338  
A:Accession: B26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADP>  
A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446  
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.  
Cell. Mol. Neurobiol. 12, 447-454, 1992  
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone  
A:Reference number: A48410; MUID:93105480

A:Accession: A48410  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-92 <MAI>  
A:Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060  
A:Experimental source: thymus  
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P.121083)  
C:Genetics:  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: stimulates pituitary secretion of luteotropin and follitropin  
C:Note: gonadoliberin-associated protein may have prolactin release inhibiting activi  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
F:1-23/Dominant: signal sequence #status predicted <SIG>  
F:24-92/Product: progonaoliberin #status predicted <PCN>  
F:24-33/Product: gonadoliberin #status predicted <GIN>  
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 87.9%; Score 51; DB 1; Length 92;  
Best Local Similarity 88.9%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 HWSYGLRPG 10  
Db 25 HWSYGLRPG 33

RESULT 8  
RHAQI  
gonadoliberin I - American alligator  
N:Alternate names: gonadotropin-releasing hormone I  
C:Species: Alligator mississippiensis (American alligator)  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Mar-1997  
C:Accession: A60066  
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan  
Regul. Pept. 33, 105-116, 1991  
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains  
A:Reference number: A60066; MUID:91352338  
A:Accession: A60066  
A:Molecule type: protein  
A:Residues: 1-10 <LOV>  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 81.0%; Score 47; DB 1; Length 10;  
Best Local Similarity 77.8%; Pred. No. 0.01;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 HWSYGLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 9  
I50644  
gonadoliberin I precursor - chicken  
N:Alternate names: gonadotropin-releasing hormone I  
C:Species: Gallus gallus (chicken)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999  
C:Accession: I50644; S33507  
R:Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.  
J. Mol. Endocrinol. 11, 19-29, 1993  
A:Title: Characterization of the chicken pregonadotropin-releasing hormone-I gene  
A:Reference number: I50644; MUID:94059355  
A:Accession: I50644  
A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-92 <DD2>  
A:Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA9246.1; PID:g311612  
C:Genetics: 47/3, 79/3  
C:Superfamily: gonadoliberin

Query Match 81.0%; Score 47; DB 2; Length 92;  
Best Local Similarity 77.8%; Pred. No. 0.089;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
||| | | |  
Db 25 HWSYGIQPG 33

RESULT 10  
150739  
gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)  
C:Species: Haplochromis burtoni  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: 150739  
R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.  
Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995  
A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles  
A:Reference number: 150739; MUID:95396797  
A:Accession: 150739  
A:Status: preliminary; translated from GH/EMBL/DBD  
A:Molecule type: mRNA  
A:Residues: 1-98 <WHI>  
A:Cross-references: EMBL:U31865; NID:g905398; PIDN:MAC59691.1; PID:g905399  
C:Superfamily: gonadoliberin

Query Match 77.6%; Score 45; DB 2; Length 98;  
Best Local Similarity 77.8%; Pred. No. 0.22;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
||| | | |  
Db 24 HWSYGLSPG 32

RESULT 11  
E64728  
yabn protein - Escherichia coli  
C:Species: Escherichia coli  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
C:Accession: E64728  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A: Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:9742617  
A:Accession: E64728  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-551 <BIAT>  
A:Cross-references: GB:AE000117; GB:U00096; NID:g1786250; PIDN:AACT3180.1; PID:g1786256;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: yabn

Query Match 72.4%; Score 42; DB 2; Length 551;  
Best Local Similarity 66.7%; Pred. No. 4.2;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
||| | | |  
Db 179 HWRFLRPG 187

RESULT 12  
G82443  
conserved hypothetical protein VCA0578 [imported] - Vibrio cholerae (group O1 strain  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 15-Sep-2000  
C:Accession: G82443  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.  
charadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers  
L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: G82443  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-565 <HEI>  
A:Cross-references: GB:AE004388; GB:AE003853; NID:g9657979; PIDN:AAF96480.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Map position: 2

Query Match 72.4%; Score 42; DB 2; Length 565;  
Best Local Similarity 66.7%; Pred. No. 4.3;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
||| | | |  
Db 179 HWRFLRPG 187

RESULT 13  
RHID15  
gonadoliberin I precursor - Sharptooth catfish  
N:Alternate names: gonadoliberin, catfish-type; gonadotropin-releasing hormone I (GNR  
N:Contents: gonadoliberin I; gonadoliberin I-associated protein form I; gonadoliberin  
C:Species: Clarias gariepinus (Sharptooth catfish)  
C>Date: 30-Sep-1993 #sequence\_revision 18-Mar-1997 #text\_change 18-Jun-1999  
C:Accession: S45602; S45601; JCI1242; S42936; S42937  
R:Bogerd, J.; Zandbergen, T.; Andersson, E.; Goos, H.  
Eur. J. Biochem. 222, 541-549, 1994  
A:Title: Isolation, characterization and expression of cDNAs encoding the catfish-type  
A:Reference number: S45600; MUID:94291651  
A:Accession: S45602  
A:Molecule type: mRNA  
A:Residues: 1-80 <BOG1>  
A:Cross-references: EMBL:X78049; NID:g459433; PIDN:CAA54971.1; PID:g459434  
A:Note: gonadoliberin I-associated protein form I  
A:Accession: S45601  
A:Molecule type: protein  
A:Residues: 1-46, 'S', 48-59, 'G', 61-80 <BOG2>  
A:Cross-references: EMBL:X78048; NID:g459433; PIDN:CAA54970.1; PID:g459432  
A:Note: gonadoliberin I-associated protein form II, presumed to be a polymorphic form  
R:Bogerd, J.; Li, K.W.; Janssen-Dommerholt, C.; Goos, H.  
Biochem. Biophys. Res. Commun. 187, 127-134, 1992  
A:Title: Two gonadotropin-releasing hormones from African catfish (Clarias gariepinus  
A:Reference number: JCI1242; MUID:92392313  
A:Accession: JCI1242  
A:Molecule type: protein  
A:Residues: 22-31 <BOG3>  
A:Experimental source: brain  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/21/Domain: signal sequence #status predicted <SIG>  
F:22-31/Product: gonadoliberin I #status experimental <MAT1>  
F:33-80/Product: gonadoliberin I-associated protein #status predicted <MAT2>  
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi  
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 70.7%; Score 41; DB 1; Length 80;

Best Local Similarity 66.7%; Pred. No. 0.95;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||:| ||  
Db 23 HWSHGLNPG 31

RESULT 14

T01937  
hypothetical protein F1104.7 - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 24-Mar-1999

C:Accession: T01937

R:Abu-Threideh, J.; Stoneking, T.; Langston, Y.; Trevaaskis, E.

submitted to the EMBL Data library, October 1998

A:Reference number: Z14466

A:Accession: T01937

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1508 <ABU>

A:Cross-references: EMBL:AF096370; NID:g3695372; PID:g3695380

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Note: F1104.7

Query Match 67.2%; Score 39; DB 2; Length 508;  
Best Local Similarity 75.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WSYXLRPG 10  
| | | |||  
Db 80 WCYSLRPG 87

RESULT 15

T08556  
hypothetical protein F27B13.190 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 22-Oct-1999

C:Accession: T08556

R:Bevan, M.; Zimmermann, W.; Gruenewald, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16442

A:Accession: T08556

A:Molecule type: DNA

A:Residues: 1828 <BEV>

A:Cross-references: EMBL:AL050352; GSPDB:GN00062; ATSP:F27B13.190

A:Experimental source: cultivar Columbia; BAC clone F27B13

C:Genetics:

A:Gene: ATSP:F27B13.190

A:Map position: 4

A:Introns: 52/2; 87/1; 146/1; 777/3

Query Match 67.2%; Score 39; DB 2; Length 828;  
Best Local Similarity 66.7%; Pred. No. 22;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||:| ||  
Db 114 HWSYFOAPG 122

Search completed: March 2, 2001, 10:55:38  
Job time: 258 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:05 ; Search time 32.03 Seconds

(without alignments)  
14.115 Million cell updates/sec

Title: US-09-306-689-2  
Perfect score: 58  
Sequence: 1 XHMSYXLRPEXXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues  
Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	51	87.9	61	GON1_SHEEP
2	51	87.9	63	GON1_MESAU
3	51	87.9	67	GON1_MACMU
4	51	87.9	89	GON1_XENIA
5	51	87.9	90	GON1_MOUSE
6	51	87.9	91	GON1_PIG
7	51	87.9	92	GON1_HUMAN
8	51	87.9	92	GON1_RAT
9	51	87.9	92	GON1_TUPGB
10	47	81.0	10	GON1_ALUMI
11	47	81.0	92	GON1_CHICK
12	45	77.6	94	GON1_HAPBU
13	45	77.6	95	GON1_PAGMA
14	45	77.6	95	GON1_SPAU
15	42	72.4	92	GON1_CAVPO
16	42	72.4	551	GON1_ECOLI
17	41	70.7	80	GON1_CLABA
18	38	65.5	10	GON3_ONCKE
19	38	65.5	14	GON3_ONCKE
20	38	65.5	74	GON3_ONCKE
21	38	65.5	82	GON3_ONCKE
22	38	65.5	82	GON3_SALSA
23	38	65.5	82	GON3_SALSA
24	38	65.5	89	GON3_PORNO
25	38	65.5	90	GON3_HAPBU
26	38	65.5	90	GON3_PAGMA
27	38	65.5	90	GON3_SPAU
28	38	65.5	94	GON3_CAVAU
29	38	65.5	94	GON3_RUTRU
30	37	63.8	10	GON3_PETMA
31	37	63.8	584	GON2_CLOCE
32	35	60.3	390	RM03_YEAST
33	35	60.3	477	GUNA_CLOTM

## ALIGNMENTS

RESULT ID	1	GON1_SHEEP	STANDARD:	PRT:	61 AA.
AC	Q28588:				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	PROGONADOLIBERIN I PRECURSOR (CONTRAINS: GONADOLIBERIN I (LHRH I) (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I) (FRAGMENT).				
DE	HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I) (FRAGMENT).				
GN	GNRH1 OR GNRH OR LHRH.				
OS	Ovis aries (Sheep).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.				
CC	[1]				
RP	SEQUENCE OF 12-61 FROM N.A.				
RC	STRAIN-WESTERN RANGE: TISSUE-HYPOTHALAMUS;				
RA	Rodriguez R.E., Wise M.E.;				
RL	Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE OF 1-10.				
RX	MEDLINE=72094314; PubMed=4550508;				
RA	Burgus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,				
RA	Fellows R., Blackwell R., Vale W., Guillemin R.;				
RT	"Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass spectrometry-decapeptide-Edman degradation).";				
RT	Proc. Natl. Acad. Sci. U.S.A. 69: 278-282(1972).				
CC	[1]				
CC	THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.				
CC	-----				
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL; U02517; AAA03433.1; -.				
DR	PIR; A93780; RSHRG.				
DR	INTERPRO; IPR002012; -.				
DR	PFAM; PF00446; GNRH; 1.				
DR	PROSITE; PS00473; GNRH; 1.				
KW	Cleavage on pair of basic residues; Hormone; Amideation; Hypothalamus; Placenta.				
FT	NON_TER	1			
FT	CHAIN	1	>61		PROGONADOLIBERIN I.
FT	PEPTIDE	1	10		GONADOLIBERIN I.
FT	PEPTIDE	14	>61		GNRH-ASSOCIATED PEPTIDE I.
FT	ACT_SITE	3	3		APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.

FT MOD\_RES 1 1 PYROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).  
FT NON\_TER 61 61  
SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 61;  
Best Local Similarity 88.9%; Pred. No. 0.0045;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 2 HWSYGLRPG 10

## RESULT 2

CON1\_MESAU STANDARD; PRT; 63 AA.

AC 009163;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
(LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]  
(FRAGMENT).  
DE GNRH I OR GNRH OR LHRH.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Mesocricetus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Jansen H.T., Stevens P.J., Zeidler P., Lehman M.N.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. IT STIMULATES  
THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
HORMONES.  
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U91938; AAB51302.1; -.  
DR INTERPRO: IPR002012; -.  
DR PFAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Placenta.  
FT NON\_TER 1 1  
FT CHAIN 1 >63 PROGNADOLIBERIN I.  
FT PEPTIDE 1 10 GONADOLIBERIN I.  
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY  
SIMILARITY).  
FT ACT\_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
ACTIVITY (BY SIMILARITY).  
FT MOD\_RES 1 1 PYROLIDONE CARBOXYLIC ACID (BY  
SIMILARITY).  
FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY  
SIMILARITY).  
FT NON\_TER 63 63  
SQ SEQUENCE 63 AA; 7370 MW; FC9499567F77180 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 63;  
Best Local Similarity 88.9%; Pred. No. 0.0046;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 2 HWSYGLRPG 10

## RESULT 3

CON1\_MACMU STANDARD; PRT; 67 AA.

AC P55247;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
(LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]  
(FRAGMENT).  
DE GNRH I OR GNRH OR LHRH.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
Cercopithecinae; Macaca.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-HYPOTHALAMUS;  
RX MEDLINE-95124501; PubMed=7545971;  
RA Y.J., Costa M.E., Ojeda S.R.;  
RT "Developmental expression of the genes encoding transforming growth  
factor alpha and its receptor in the hypothalamus of female rhesus  
macaques";  
RL Neuroendocrinology 60:346-359(1994).  
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. IT STIMULATES  
THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
HORMONES.  
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: S75918; AAB33096.1; -.  
DR INTERPRO: IPR002012; -.  
DR PFAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 5 BY SIMILARITY.  
FT CHAIN 6 >67 PROGNADOLIBERIN I.  
FT PEPTIDE 6 15 GONADOLIBERIN I.  
FT PEPTIDE 19 >67 GNRH-ASSOCIATED PEPTIDE I.  
FT ACT\_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
ACTIVITY (BY SIMILARITY).  
FT MOD\_RES 6 6 PYROLIDONE CARBOXYLIC ACID (BY  
SIMILARITY).  
FT MOD\_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY  
SIMILARITY).  
FT NON\_TER 67 67  
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 67;  
Best Local Similarity 88.9%; Pred. No. 0.0049;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 7 HWSYGLRPG 15

```

RESULT      4
GONL_XENLA  STANDARD:      PRT:      89 AA.
AC P45656;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
  (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
  [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FOREBRAIN;
RC MEDLINE=94185563; PubMed=8137750;
  RA Hayes W.P., Gray S., Battey J.F.;
  RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
  RT mammalian like expression pattern and conserved domains in
  RT GNRH-associated peptide, but brain onset is delayed until
  RT metamorphosis.";
  RL Endocrinology 134:1835-1844(1994).
  CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
  CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
  CC -----
  CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL; L28040; AAA49728.1; -.
  DR INTERPRO: IPR002012; -.
  DR PRAM; PF00446; GNRH.1.
  DR PROSITE; PS00473; GNRH.1.
  KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
  KW signal.
  FT SIGNAL. 1 23
  FT CHAIN 24 89
  FT PEPTIDE 24 33
  FT CHAIN 37 89
  FT PEPTIDE 37 85
  FT MOD_RES 24 24
  FT MOD_RES 33 33
  SO SEQUENCE 89 AA; 10246 MW; 6FAF36FBAE0D4284 CRC64;
Query Match 87.9%; Score 51; DB 1; Length 89;
Best Local Similarity 88.9%; Pred. No. 0.0065;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 HWSYTLRPG 10
  |||||
Db 25 HWSYTLRPG 33
RESULT      5
GONL_MOUSE  STANDARD:      PRT:      90 AA.
AC P13562;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (LHRH I)
  (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
  DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
  DE I).
  GN GNRH1 OR GNRH.
  OS Mus musculus (Mouse).

```

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Euteheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87069928; PubMed=302417;
RA Mason A.J., Haylick J.S., Zoeller R.T., Young W.S. III,
RA Phillips H.S., Nikolic K., Seeburg P.H.;
RT "A deletion truncating the gonadotropin-releasing hormone gene is
RL responsible for hypogonadism in the hpg mouse."
RL Science 234:1366-1371(1986).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/sib.ch).
CC CC send an email to license@sib.ch).
CC -----
DR EMBL; M14872; AAA37717.1; -.
DR MGD; MGI:95789; GnRH.
DR INTERPRO; IPR002012; -.
DR PFAM; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
KM SIGNAL
FT 1 21
FT CHAIN 22 90 PROGNADOLIBERIN I.
FT PEPTIDE 22 31 GONADOLIBERIN I.
FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 22 22 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
FT SEQUENCE 90 AA; 10337 MW; 1C0766F4826E4D9 CRC64;
OY 2 HWSYLRPG 10
DY |||||
DB 23 HWSYLRPG 31

```

RT "Structure of the porcine LH- and FSH-releasing hormone. II.  
RT Confirmation of the proposed structure by conventional sequential  
RT analyses.";  
RL Biochem. Biophys. Res. Commun. 44:459-463(1971).  
RN [3]  
RP SYNTHESIS OF GONADOLIBERIN.  
RX MEDLINE=72065376; PubMed=4942726;  
RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;  
RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-  
RT phase method.";  
RL Biochem. Biophys. Res. Commun. 45:822-827(1971).  
RN [4]  
RP SYNTHESIS OF GONADOLIBERIN.  
RX MEDLINE=72117544; PubMed=4946275;  
RA Baba Y., Arimura A., Schally A.V.;  
RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";  
RL Biochem. Biophys. Res. Commun. 45:483-487(1971).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L32864; AAA31066.1; -;  
CC PIR: A01411; RHHGG.  
CC DR INTERPRO: IPR002012; -;  
CC PFAM: PF00446; GNRH: 1.  
CC PROSITE: PS00473; GNRH: 1.  
CC DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
CC Placenta; Signal.  
CC FT SIGNAL 1 23  
CC FT CHAIN 1 23  
CC FT PEPTIDE 24 33  
CC FT PEPTIDE 34 91  
CC FT ACT\_SITE 26 26  
CC MOD.RES 24 24  
CC MOD.RES 33 33  
CC SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;  
OY 2 HWSYXLRPG 10  
DB 25 HWSYXLRPG 33  
RESULT 7  
GON1\_HUMAN STANDARD: PRT; 92 AA.  
AC P01148;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)  
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
DE HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED  
DE PEPTIDE I].  
DE GNRH1 OR GNRH OR LHRH.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=8936682; PubMed=2671939;  
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;  
RT "The complete nucleotide sequence of the human gonadotropin-releasing  
RT hormone gene.";  
RL Nucleic Acids Res. 17:6403-6403(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86094338; PubMed=2867548;  
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;  
RT "Isolation of the gene and hypothalamic cDNA for the common precursor  
RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
RT factor in human and rat.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85012739; PubMed=6090951;  
RA Seeburg P.H., Adelman J.P.;  
RT "Characterization of cDNA for precursor of human luteinizing hormone  
RT releasing hormone.";  
RL Nature 311:666-668(1984).  
RN [4]  
RP SEQUENCE OF 24-33.  
RX MEDLINE=83126573; PubMed=6760865;  
RA Tan L., Rousseau P.;  
RT "The chemical identity of the immunoreactive LHRH-like peptide  
RT biosynthesized in the human placenta.";  
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS),  
CC LUTERPUSE OR LUTERLEF (FERRING PHARMACEUTICALS) AND RELISORM  
CC (SERONO).  
CC -----  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X01059; CAA25526.1; -;  
CC EMBL: M12578; AAA35916.1; -;  
CC DR EMBL: X15215; CAA33285.1; -;  
CC PIR: A01410; RHHGG.  
CC DR PIR: A26173; A26173.  
CC DR PIR: S05308; S05308.  
CC DR MIM: 152760; -;  
CC DR INTERPRO: IPR002012; -;  
CC PFAM: PF00446; GNRH: 1.  
CC PROSITE: PS00473; GNRH: 1.  
CC DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
CC Placenta; Pharmaceutical; Signal.  
CC FT SIGNAL 1 23  
CC FT CHAIN 1 23  
CC FT PEPTIDE 24 33  
CC FT PEPTIDE 37 92  
CC FT ACT\_SITE 26 26  
CC MOD.RES 24 24  
CC MOD.RES 33 33  
CC CONFLICT 16 16  
CC SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;  
OY 2 HWSYXLRPG 10  
Query Match 87.9%; Score 51; DB 1; Length 92;  
Best Local Similarity 88.9%; Pred. No. 0.0067; 1; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 1;

```

DB 25 HWSYGLRPG 33

|||||
RESULT 8
GONL_RAT STANDARD: PRT: 92 AA.
AC P07490:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
DE I].
GN GNRI OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86094338; PubMed-2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-89384661; PubMed-2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
RT hypothalamic expression.";
RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RX MEDLINE-93105480; PubMed-1468115;
RA Walter C.C., Marchetti B., Lebeuf R.D., Bialock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE=HEART;
RX MEDLINE-87149087; PubMed-3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus.";
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, S50870; AAB24572.1; -
DR EMBL, M12579; AAA41263.1; -
DR EMBL, M31670; AAA41264.1; -
DR EMBL, M15527; AAA42141.1; ALT_SEQ.
DR EMBL, M15529; AAA42139.1; -
DR EMBL, M15528; -; NOT_ANNOTATED_CDS.
DR PIR, B26173; RHRTG.
DR PIR, A48410; A48410.
DR INTERPRO; IPR002012; -

```

```

DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 92;
Best Local Similarity 88.9%; Pred. No. 0.0067;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33

RESULT 9
GONL_TUPGB STANDARD: PRT: 92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRI OR GNRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupia.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HYPOTHALAMUS;
RX MEDLINE-97079639; PubMed-8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U63326; AAB16837.1; -
DR INTERPRO; IPR002012; -
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT BY SIMILARITY:
FT PROGNADOLIBERIN I.
FT GONADOLIBERIN I.
FT GNRH-ASSOCIATED PEPTIDE I.
FT APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.

```

```

FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 33 33 SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
SIMILARITY).

Query Match 87.9%; Score 51; DB 1; Length 92;
Best Local Similarity 88.9%; Pred. No. 0.0067;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
    |||||
Db 25 HWSYGLRPG 33

RESULT 10
GONI_ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I)
DE (LULIBERIN I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
RN [1]
RP SEQUENCE.
RC TISSUE-BRAIN;
RC MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
RT from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR INTERPRO: IPR002012;
DR PIR: A60066; RHA01.
DR INTERPRO: IPR002012;
DR PIR: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
DR PIR: PF00446; GNRH; 1.
DR PIR: PF00446; GNRH; 1.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 81.0%; Score 47; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 0.0042;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
    |||||
Db 2 HWSYGLRPG 10

RESULT 11
GONI_CHICK STANDARD; PRT; 92 AA.
AC P37042; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RN [1],

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEGHORN;
RC MEDLINE=94059355; PubMed=7902095;
RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
RT "Characterization of the chicken preprogonadotropin-releasing
RT hormone-I gene.";
RL J. Mol. Endocrinol. 11:19-29(1993).
RN [2]
RN SEQUENCE OF 24-33.
RC TISSUE-HYPOTHALAMUS;
RC MEDLINE=82265778; PubMed=7050119;
RX King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RT hormone. II. Isolation and characterization.";
RL J. Biol. Chem. 257:10729-10732(1982).
RN [3]
RN SEQUENCE OF 24-33.
RC TISSUE-HYPOTHALAMUS;
RA King J.A., Millar R.P.;
RT "Structure of avian hypothalamic gonadotropin-releasing hormone.";
RL S. Afr. J. Sci. 78:124-125(1982).
RN [4]
RN SYNTHESIS OF 24-33.
RX MEDLINE=82265777; PubMed=7050118;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RT hormone. I. Structural determination on partially purified
RT material.";
RL J. Biol. Chem. 257:10722-10728(1982).
CC CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69491; CAA49246.1;
DR PIR: S33507; S33507.
DR INTERPRO: IPR002012;
DR PIR: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
DR PIR: PF00446; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT MOD_RES 24 24 GNRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 92 AA; 10206 MW; 61AEB7EBAF508B6A CRC64;

Query Match 81.0%; Score 47; DB 1; Length 92;
Best Local Similarity 77.8%; Pred. No. 0.036;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
    |||||
Db 25 HWSYGLRPG 33

RESULT 12
GONI_HAPBU STANDARD; PRT; 94 AA.
AC P51918; 093387;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)

```

DE (LH-RH I) (LULIBERIN I).  
GN GNRH1.  
OS Haplochromis burtoni.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;  
OC Cichlidae; Astatotilapia.  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=95396797; PubMed=7667296;  
RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;  
RT "Three gonadotropin-releasing hormone genes in one organism suggest  
RT novel roles for an ancient peptide."  
RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).  
[2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=99061842; PubMed=98043638;  
RA White R.B., Fernald R.D.;  
RT "Ontogeny of gonadotropin-releasing hormone (GnRH) gene expression  
RT reveals a distinct origin for GnRH-containing neurons in the  
RT midbrain."  
RL Gen. Comp. Endocrinol. 112:322-329(1998).  
[3]  
RN SEQUENCE OF 23-32.  
RP TISSUE=PIUITARY;  
RC MEDLINE=95372591; PubMed=7644702;  
RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,  
RA White S.A., Francis R.C., Fernald R.D., Licht P., Warty C.,  
RA Sherwood N.M.;  
RT "Primary structure of solitary form of gonadotropin-releasing hormone  
RT (GnRH) in cichlid pituitary: three forms of GnRH in brain of cichlid  
RT and pumpkinseed fish."  
RL Regul. Pept. 57:43-53(1995).  
-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE  
RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALMIC-PITUITARY-  
GONADAL AXIS.  
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS  
TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOTHYSEAL AXONS.  
CC -1- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
DR EMBL: U31865; AAC59691.1; -  
DR EMBL: AF076961; AAC27716.1; -  
DR INTERPRO: IPR002012; -  
DR PIRAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Signal; Multigene family.  
FT SIGNAL 1 22  
FT CHAIN 23 94  
FT PEPTIDE 23 94  
FT MOD\_RES 36 94  
FT MOD\_RES 23 23  
FT MOD\_RES 32 32  
FT CONFLICT 86 94  
FT SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;  
Query Match 77.6%; Score 45; DB 1; Length 94;  
Best Local Similarity 77.8%; Pred. No. 0.085;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 13  
GONL\_PAGMA STANDARD; PRT; 95 AA.  
AC P70074;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)  
DE (LH-RH I) (LULIBERIN I).  
OS Pagrus major (Red sea bream) (Chrysophrys major).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;  
OC Sparidae; Chrysophrys.  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=BRAIN;  
RC Okuzawa K., Granneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;  
RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY  
CC SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
DR EMBL: D86582; BAA13129.1; -  
DR INTERPRO: IPR002012; -  
DR PIRAM: PF00446; GNRH; 1.  
DR PROSITE: PS00473; GNRH; 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Signal; Multigene family.  
FT SIGNAL 1 23  
FT CHAIN 24 95  
FT PEPTIDE 24 95  
FT MOD\_RES 37 95  
FT MOD\_RES 24 24  
FT MOD\_RES 33 33  
FT SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;  
Query Match 77.6%; Score 45; DB 1; Length 95;  
Best Local Similarity 77.8%; Pred. No. 0.085;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA MEDLINE=95268499; PubMed-7749463;
RX Gottlieb Y., Elitzur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-
RL releasing hormone from the gilthead seabream (Sparus aurata).";
RM Mol. Mar. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE-BRAIN;
RA MEDLINE=95083645; PubMed=7991588;
RX Powell J.F.F., Zohar Y., Elitzur A., Park M., Fischer W.H.,
RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
RL brains of one species.";
RM Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U30320; AAA75469.1; -
DR INTERPRO: IPR002012; -
DR PFM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT CHAIN 1 25
FT PEPTIDE 26 35
FT MOD_RES 39 95
FT MOD_RES 26 26
FT MOD_RES 35 35
SQ SEQUENCE 95 AA; 10753 MW; 49313FD6D6B87DA CRC64;

Query Match 77.6%; Score 45; DB 1; Length 95;
Best Local Similarity 77.8%; Pred. No. 0.085;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
DB 27 HWSYGLSPG 35

RESULT 15
GONI_CAVPO STANDARD; PRT; 92 AA.
ID GONI_CAVPO
AC 054713;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE I) (GNRH I) (LULIBERIN I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRH1 OR GNRH OR LHRH.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eultheria; Rodentia; Hystriognathii; Cavidae; Cavia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HARTLEY WHITE; TISSUE=HYPOPHALAMUS;
RX MEDLINE=97462693; PubMed=9322920;
RA Jimenez-Linan M., Rubin B.S., King J.C.;
RT "Examination of guinea pig luteinizing hormone-releasing hormone gene

```

```

RT reveals a unique decapeptide and existence of two transcripts in the
RT brain.";
RL Endocrinology 138:4123-4130(1997).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF033346; AAB87688.1; -
DR INTERPRO: IPR002012; -
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT CHAIN 1 23
FT PEPTIDE 24 92
FT PEPTIDE 24 33
FT ACT_SITE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10279 MW; ACF74613F456D663 CRC64;

Query Match 72.4%; Score 42; DB 1; Length 92;
Best Local Similarity 66.7%; Pred. No. 0.29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
DB 25 YWSYGVVRPG 33

```

Search completed: March 2, 2001, 11:01:19  
Job time: 434 sec



4

**About:** Results were produced by the Gencore software, version 4.5  
Copyright (c) 1993-2000 Compugen Ltd.

```
-Model=Irfame+_n2p_model -DEV=xlp
-Q=Cgcn2.1/USPO2.spo1/USO3.9206689/runat.02032001-102826.9576/app-query.fasta_1.2389
-DB=Pending_Patents_AA -QFmr=fastan -SUFFIX=rap -GAPOP=12.000
```

```

-GEAPEAT=4.000 -MINMATCH=0.050 LOOPCL=0.000 -LOOPEXT=0.000
-OGAPEXT=4.500 -OGAPEXTC=0.100 -XGAPD=10.000 -XGAPEXT=0.500
-DEGAPD=6.000 -FGAPEXT=7.000 -XGAPD=10.000 -YGAPEXT=0.500
-FLOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bloms62
-TRANS-human40.cdi -LIST=45 -DOCALL=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pts
-NORM=0 -MILEN=0 -MAXILEN=2000000000
-USER-US09306689.ecgcn1_1_839 -MCPD=6 -ICPU=3 -LONGLOG -NO_XLPHYX
-WALT -THREADS=1

```

```
Query: 05-09-306-689-12
Query length: 2088
```

Database: Pending\_Patents\_AA:  
Database sequences: 957798

Search time (sec): 459.970000

Sequence	Strd Orig	ZScore	EScore Len	Documentation
----------	-----------	--------	------------	---------------

[illegible]

seq. documentation block:  
 : Sequence 47, Application US/08160882  
 : GENERAL INFORMATION:  
 : APPLICANT: Russel-Jones, Gregory J.  
 : APPLICANT: Stewart, Andrew G.  
 : APPLICANT: Tsonis, Con G.  
 : TITLE OF INVENTION: TriaT/Neuropeptide-Y Fusion Proteins  
 : NUMBER OF SEQUENCES: 55  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Foley & Lardner  
 : STREET: 3000 K Street, N.W., Suite 500  
 : CITY: Washington, D.C.  
 : COUNTRY: USA  
 : ZIP: 20007-5109  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.25  
 : CURRENT APPLICATION NUMBER: US/08/160,882  
 : FILING DATE: 03-DEC-1993

	alignment_scores:	
Quality:	368.00	Length: 1017
Ratio:	4.779	Gaps: 7
Percent Similarity:	76.238	Percent Identity: 68.311

Align seg 1/1 to: US-08-160-882-47 from: 1 to: 84

```

22  TGTGACAAATTGGAGCTAGAGGCTGGGCCCTGGACAGCGTCTCTCAAGATTG  71
    |||||
2  SerGluHstIrrpSerTyrGlyLeuAtrProGly.....GluHstIrr  15

72  GAGCTACGAGCGCGCTCCGGGTGGCTGTACGACAGCATGAGATACGAGGC  122
    |||||
15  pSerYrGlyLeuAtrProGly.....GluHstIrrpSerTyrGlyL  29

122  TGGCGCCCTGGAGCGGCTAGCCAAATTGGAGCTACGAGGCTCGTCCGGGT  171
    |||||
29  euAtrpProGly.....GluHstIrrpSerTyrGlyLeuAtrProGly  42

```

```

172  GGATGCTACGACATTTGGAGCTACAGCGCCTGGCCCTGGACAGCGTCTTCACAGA 222
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
43   .....GluHisTrpSerTyrGlyLeuArgProGly.....GluHis 54

222  TTGGAGCTACAGCGCCTGGCCTGGGGGTGGCTACGACAGATTGAGATACG 27
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
54  strPserTyrGlyLeuArgProGly.....GluHisTrpSerTyrG 68

272  GCCTGCGCCCTGAGCGACGCGGTACCCACAGATTGAGCTACAGCGCTGCTCG 322
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
68  lyLeuArgProGly.....GluHisTrpSerTyrGlyLeuArgPro 81

```

322	GGT	32
82	GLY	82

seq\_name: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:US-09-506-078-27

```

seq_documentation_block:
; Sequence 27, Application US/09506078
; GENERAL INFORMATION:
; APPLICANT: Pfizer Products Inc
; TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A
; TITLE OF INVENTION: DUAL IMMUNE RESPONSE

```

FILE REFERENCE: PC10202A  
CURRENT APPLICATION NUMBER: US/09/506,078  
CURRENT FILING DATE: 2000-02-16  
EARLIER APPLICATION NUMBER: N/A  
EARLIER FILING DATE: 1999-02-17  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 27  
LENGTH: 442  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:  
US-09-506-078-27  
4GnRH-LmGD-4GnRH encoded by pOE-GnRH:gd:GnRH

alignment\_scores:  
Quality: 304.50 Length: 634  
Ratio: 1.208 Gaps: 25  
Percent Similarity: 39.748 Percent Identity: 22.082

## alignment\_block:

US-09-306-689-12 x US-09-506-078-27

Align seg 1/1 to: US-09-506-078-27 from: 1 to: 442

```
178 CAGATTGGAGCTACGGCTCGCCCTGGCAGCGGTTCTCAAGTTGAG 227
    ::::|||||:
19 GluHisTrpSerTyrGlyLeuArgProGly.....GluHisTrpSe 32
228 CTAGAGCGCTCGGCTCGGCTGAGCTAGCCAGTGGAGCTAGCGGCTGC 277
    |||||:|||||:
32 rTyrGlyLeuArgProGly.....GluHisTrpSerTyrGlyLeuA 46
278 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTCGCCGGGTGGA 327
    |||||:
46 rGProGly.....GluHisTrpSerTyrGlyLeuArgProGlySer 59
328 TCTAGCTTCCCAAAACTGGGGCAAAAAATTTCCCTATATCC... 375
    |||||:
60 MetSerLeuPro...ThrProAlaProArgValThrValTyrValAsp 75
376 .....CAAATTACCAATATGATGATGTCGAAACAGGTAATG 409
    ::::|||||:
75 oProAlaTyrProMetProArgTyrAsnTyr...ThrGluArgTrpHisT 91
410 GTTACAGAGATTAGTCAAAGCG.....GCCGAAGAGTTGGGGATT 450
    ::::|||||:
91 hrThnGlyProIleProSerProPheAlaAspGlyArgGluGlnProVal 107
451 GAGGTACAAAGAGAAGAACGCAATATATTGCCAACAGCTCAACACAGTT 500
    |||||:
108 GluValArg..... 110
501 AGGCACGATTCAAACCGCTATTGGCTTAACAGCGTGCATTTGTAT 550
    |||||:
111 .....TyrIleThrSerIleAlaIleAlaCysAspMetLeuAlaLeuIle 125
551 CCGGTCCACAAATGATTAATTTGCTACGAAACAACTAAGAGCCACACA 600
    ::::|||||:
125 LaAspProGlnVal.....GlyArgThr 132
601 TTAGTCTTCGCCGAAGCATGTGTACAAATGCAATTAAGCCAAACGCT 650
    |||||:
133 LeuTrpGluAlaValArgArgHisAlaArgValArgHisAlaThrValI 149
651 ATTATCTGCGCATTCATCTATTAGGCTCAGTATTGGCTGGAATGAT 700
    |||||:
149 eTrpTyrIleSLeuGluSerGlyCysAlaArgProLeuTyrTrpMetGlu 166
701 TAGATGAGCGCTTACAGATAACAGCAACACACATGCTCTTGTAAAGCT 750
    |||||:
```

```
166 yTrhGluCysGluProArgLysHisPheGlyTyrCysArgTyrArgThr 182
751 GGCTTGAGGTCACAAATTCATTATGAAATATTCCTATTCAGAAA 800
    ::::|||||:
183 Pro...ProPheTrpAspSerPheLeuAlaGlyPheAla...TyrP 196
801 AACACTTGACGAATTTGTGAGCAAAATTAAGTCAATTTGTTCAAAACTAC 850
    |||||:
196 oThrAspAspGluLeuGly..... 202
851 AAAATATCAAAAGCTTAGGACCTTAGAGACAAACATCAAAATATCGGT 900
202 ..... 202
901 GGAATTGATAAGCTGGCCTTGTGTTAGATGATATCTCAGGCGATATTATC 950
202 ..... 202
951 GGGCCAAACAGCTGCACTTGACTTGCGAGATAAATGCTTCAACAGCTA 1000
    |||||:
203 .....LeuIleMet.....AlaIleProAla 210
1001 AAAAAGTGGGTGGCGGTTTGAATTCGCAAAACCAAGTTGTTGTAATTT 1050
    ::::|||||:
210 rGLeuValGluGlyGlnTyrArgArgAlaLeuTyrIleAspGly..... 224
1051 ACCAAGCGCTTCTCTCTACATTTTAGCCCAACGCTTGACAGAGCTTT 1100
    |||||:
225 ThrValAlaTyrThrAspPheMetVal..... 233
1101 ATCTCAACTGGGCGCTGGCTGCTTAATGCTTCTACTGTTCTCTTG 1150
233 ..... 233
1151 CGATTAGCCCTTAGCATTTGCCGTAATGCCGTAATTAATCAATGCA 1200
233 ..... 233
1201 AAAAGTTAGAGATTATGCCGAAACGCTTAAATTAAGCTATGACGG 1250
    |||||:
234 ...SerLeuProAlaGlyAspCysTrpPheSerLysLeuGly..... 246
1251 AGATTAATTATTACAGAAATATACAGCGGGGAAACAGGACTATTGATCAT 1300
246 ..... 246
1301 CGTTTACTGCATTAATACCGCATTTGCCGCTATTGCTGGTGCTGCT 1350
    |||||:
247 .....AlaIleArgGlyTyrTrpPheGlyAlaCys 256
1351 GCTGTCGAGCGGATTTAACATTTGAAAGGTTAAACATATATCTGTGCAT 1400
    |||||:
257 PheProAlaArgAspTyrGluGlnLysValLeuArgLeuTrpTyrLe 273
1401 CACGAAT.....AGCAAAAAAGAAAGTGACCATTCAAAACCTGCT 1441
    |||||:
273 uThrGlnTyrTrpGlnGluAlaHisLysAlaIleValAspTyrTrp 290
1442 TCCGAGAG.....GCTGATTTTGTCTAAAGAACTCCTAAT 1476
    |||||:
290 hMetArgHisGlyGlyValValProProTyrPheGluGluSerLysGly 306
1477 TATAAGCAATAA.....GA 1493
307 TyrGluProProProAlaIleAspGlySerProAlaProProGlyAs 323
1494 TGAGAAATCGAAGAAATCATCGCTCAAAATGGCAGCGGATCACTCAA 1543
    ::::|||||:
323 pAspGluAlaArgGluAspGlyGlnTyrThrGluAspGlyValAlaIleGly 340
1544 AGCAAGTTGATGATCTTATCGCAAAAGGTAAACGCAAAATTAACCAAGAT 1593
    ::::|||||:
340 rGlu.....GlyAsnGlyGlyProProGlyPro 349
```

```

1594 GACGATCAAAAGTTGTGATTAATGAAATGCTCAACATAGCAAAA 1643
    |||
350 GLU..... 350
1644 TGTGACAAACAGCTTAGATTAATCTCATCTGTAAGTCATTACCT 1693
    |||
351 .....GlyaspGlyIuserGlnThrProGluAlaAsnGly 363
1694 CGCTAATATGTTGAGAAATGATTAAGTGGCTCCAACTTAATGTTGGAT 1743
    :
363 LysAlaGluGlyProGlyProGlyProSerProAspAlaAspArgPro 379
1744 CAAGTTTATCTCTCTCAATT..... 1767
    :
380 GlnGlyThrProSerLeuGlnAlaIleThrHisProProAlaProAl 396
1768 .....GCTAGGAGATCTCAGCATTTGAGCTAGCGCTGCGCTGGCA 1810
    |||||
396 arThrProAlaArgAlaProGlnHisThrSerTyrGlyLeuAlaArgProGly 412
1811 GCGGTTCTCAAGATTGAGCTAGCGCTGCGCTGCGCTGAGCTAGCCAG 1860
    :
413 .....Glu 413
1861 CATTTGAGCTAGCGCTGCGCTGCGCTGAGCGGTAAGCAAGATTGAGCTA 1910
    |||||
414 HisThrSerTyrGlyLeuAlaArgProGly.....GlnHisThrSerTyr 427
1911 CGGCTGCGCTGCGCTGAGATCTCAGCATTTGAGCTAGCGCTGCGCTG 1960
    |||||
427 rGlyLeuAlaArgProGly.....GlnHisThrSerTyrGlyLeuAlaArgProG 442
1961 GC 1962
    ||
442 Ly 442

```

seq\_name: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:US-60-120-454-27

```

seq_documentation_block:
; Sequence 27, Application US/60120454
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Yule, Terecita D
; APPLICANT: Martind, Serge
; TITLE OF INVENTION: DUAL IMMUNE RESPONSE
; FILE REFERENCE: PCI10202
; CURRENT APPLICATION NUMBER: US/60/120,454
; EARLIER APPLICATION NUMBER: N/A
; EARLIER FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 27
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: 46nH-tmgD-46nH encoded by pQE-gnH:gd:gnH
US-60-120-454-27

```

```

alignment_scores:
    Quality: 304.50      Length: 634
    Ratio: 1.208        Gaps: 25
    Percent Similarity: 39.748    Percent Identity: 22.082

```

alignment\_block:  
US-09-306-689-12 x US-60-120-454-27 ..

```

Align seg 1/1 to: US-60-120-454-27 from: 1 to: 442
178 CAGCATTTGAGCTAGCGCTGCGCTGCGGCTTCTCAAGATTGAG 227
    :
19 GlnHisThrSerTyrGlyLeuAlaArgProGly.....GlnHisThrSe 32
228 CTAGGCGCTGCGCTGCGGCTGAGCTAGCGCTGAGCTAGCGCTGCGCTG 277
    :
32 rTyrGlyLeuAlaArgProGly.....GlnHisThrSerTyrGlyLeuAla 46
278 GCGCTGCGAGCGCTAGCCAAAGATTGAGCTAGCGCTGCGCTGCGCTGGA 327
    |||||
46 rProGly.....GlnHisThrSerTyrGlyLeuAlaArgProGlySer 59
328 TCTAGCTTCCCAAAACCTGGGCAAAAATTATCTGTATATTCCTCC... 375
    |||||
60 MetSerLeuPro...ThrProAlaProAlaValThrValTyrValAsp 75
376 .....CAAAATTACCAATATGATACTGAAACAAGTAATG 409
    :
75 oProAlaTyrProMetProArgTyrAsnTyr...ThrGluAlaGlyrHisT 91
410 GTTTACAGGATTTAGTCAAGCG.....GCCGAAAGTTGGGATTT 450
    :
91 hrThrGlyProIleProSerProPhaAlaAspGlyArgGluGlnProVal 107
451 GAGGTACAAAGAGAGACGACATTAATATTCGACACAGCTCAAAACAGTTT 500
    |||||
108 GluValAlaArg..... 110
501 AGGCAGAGATTCAAAACCGCTATTGGCTTAAGTACGCTGAGCTATGCTTAT 550
    :
111 .....TyrAlaThrSerAlaAlaIleCysAspMetLeuAlaLeuIleA 125
551 CGGCTCCACAAATTTGATAATTGCTACAGAAAACSTAACAGCGCAAGCA 600
    :
125 LaspProGlnVal.....GlyArgThr 132
601 TTAGGTTCTGCCGAAAGCATTTGTACAAATGCAAAATTAAGCCAAACCTG 650
    |||
133 LeuTrpGluAlaValAlaArgHisAlaAlaArgAlaTyrAsnAlaThrValI 149
651 ATTATGTGCATTCATATTTAGCTCAGATTTGGCTGGAATGGAT 700
    |||||
149 eTrpTyrLysIleGlySerGlyCysAlaArgProLeuTyrTyrMetGlnT 166
701 TAGATGAGGCTTACAGATAACAGCAACCAACATGCTTGTCTAAAGCT 750
    |||||
166 yThrGluGlySerProAlaGlyHisPheGlyTyrCysArgTyrArgThr 182
751 GCGTTGAGCTAACAAATTCATTAATTTGAATAATTTGCTAATTCAGTAA 800
    :
183 Pro...ProPheTrpAspSerPheLeuAlaGlyPheAla.....TyrPr 196
801 AACACTTGACGAATTTGCTGACGCAAAATTAGTCAATTTGGTCAAAACTAC 850
    |||
196 oThrAspAspGluLeuGly..... 202
851 AAAATATCAAGGCTTAGGACTTAGAGACAAACTCAAAATATTCGCT 900
202 ..... 202
901 GGACTTGATTAAGCTGGCGCTGTTAGATGTTATCTCAGGCTATTATC 950
202 ..... 202
951 GGGCGCAACAGCTGCATCTGTACTTGACAGATAAAATGTTCAACAGCTA 1000
    |||||
203 .....LeuIleMet.....AlaAlaProAlaA 210
1001 AAAAAGTGGGTGCGGGTTTGAATTTGCAAAACCAAGTTGTTGTAATATT 1050
    :
210 rGluValGluGlyGlnTyrArgArgAlaLeuTyrIleAspGly..... 224

```

```

1051 ACCAAGCCGTTCTCTTACATTTTACCCCAAGCTGTTCAGCAGGTTT 1100
    ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
225 ThrValAlaTyrThrAspHemElVal..... 233
1101 ATCTTCAACTGGCCTGTGGCTTATTCCTTACTCTTCTCTCTG 1150
233 ..... 233
1151 CGATTAGCCCATTTAGCATTTGCGGTATTGCGATAATTAATCATGCA 1200
233 ..... 233
1201 AAAAGTTTAGAGATTATGCGAAGCTTTAAAAATTAGGCTATGACGG 1250
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
234 ..SerLeuProAlaGlyAspCysTyrPheSerLysLeuGly..... 246
1251 AGATATTATTATACGAATATCAGCGGGGACAGGACTATTGATGAT 1300
246 ..... 246
1301 CGTTACTGCAATTAAATACGCGCTATGCTGCTGTGTCTCT 1350
247 ..... 247
1351 GCTGTCGAGCCGATTAAACATTTGAAAAAGTTAAACATACTTGTGAT 1400
    ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
257 PheProAlaArgAspTyrGluGlnLysValLeuArgLeuThrTyrLe 273
1401 CACGAT.....AGCAAAAAAGAGAAAGTGCATTTAAACATGCT 1441
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
273 uThGlnTyrTyrProGlnGlnAlaHisLysAla1LeValAspTyrTrp 290
1442 TCCGAGAG.....GCTGATTTGCTAAAGAGTGCCTAT 1476
    || : : : : : : : : : : : : : : : : : : : : : : : :
290 hemetaArgHisGlyGlyValValProPyrOtyrPheGlnGlnSerLysGly 306
1477 TATAAGCAACTAAA..... 1493
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
307 TyrGluProProProAlaAlaAspGlyLysSerProAlaProProGlyAs 323
1494 TGAGAAAATCAAGAAATCATCGGTCAAAATGCGAGGCGGATCAGCTCA 1543
    : : : : : : : : : : : : : : : : : : : : : : : : : :
323 pasPgluAlaArgLysAspGlnGlyGluThrGlnAspGlyAlaAlaGlyA 340
1544 AGCAAGTTGATGATCTTATCGCAAAAGTAAGCGCAAAATTAACCAAGAT 1593
    : : : : : : : : : : : : : : : : : : : : : : : : : :
340 rGglL.....GlyAsnGlyGlyProProGlyPro 349
1594 GAGCTATCAAAAGTTGTTGATTAAGTAAATGCTCAAAACATAGCAAAA 1643
    || : : : : : : : : : : : : : : : : : : : : : : : :
350 Glu..... 350
1644 TGTGACAAACAGCTTAGTAAGTAAATCTCATCTGTAAGTCATTTACT 1693
    || : : : : : : : : : : : : : : : : : : : : : : : :
351 .....GlyAspGlyGlnSerGlnThrProGlnAlaAsnGly 363
1694 CGTCTAATGATTCGAGAATGTATTAGTGCCTCAACTCAATCAATGTTGAT 1743
    : : : : : : : : : : : : : : : : : : : : : : : : : :
363 lylAlaGlnGlyGlnProLysProGlyProSerProAlaAlaAspArgPro 379
1744 CAAAGTTATCTCTCTCAATT..... 1767
    : : : : : : : : : : : : : : : : : : : : : : : : : :
380 GluGlyTrpProSerLeuGlnAla1leThrHisProProProAlaProAl 396
1768 .....GCTAGGGGATCTCAGCATGTGAGCTAGGCGCTGCGCCCTGGCA 1810
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
396 aThProAlaArgAlaProGlnHisTyrSerTyrGlyLeuArgProGly. 412
1811 GCGGTTTCAGATTTGAGCTAGGCGCTGCGGCGGCTGAGCTAGCCAG 1860
    || : : : : : : : : : : : : : : : : : : : : : : : :
113 .....Glu 413

```

```

1961 CATTGAGACTAGGCGCTGCGCCCTGGCAGCGGTAGCCAGATTGGACCTA 1910
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
414 HisTyrSerTyrGlyLeuArgProGly.....GlnHisTyrSerTyr 427
1911 CGGCTGCGTCCGGGTGATCTCAGCATTTGAGCTAGGCGCTGCGCCCTG 1960
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
427 rGlyLeuArgProGly.....GlnHisTyrSerTyrGlyLeuArgProG 442
1961 GC 1962
    ||
442 Ly 442

```

seq\_name: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:US-09-019-010-4

seq\_documentation\_block:

; Sequence 4, Application US/09019010

; GENERAL INFORMATION:

; APPLICANT: HARLAND, RICHARD

; APPLICANT: MANN, JOHN G.

; APPLICANT: ACRES, STEPHEN D.

; TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS

; MOLECULES

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROBINS & ASSOCIATES

; STREET: 90 MIDDLEFIELD ROAD, SUITE 200

; CITY: MENLO PARK

; STATE: CA

; COUNTRY: USA

; ZIP: 94025

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/019,010

; FILING DATE: 05-FEB-1998

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/036,883

; FILING DATE: 05-FEB-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: MCCracken, THOMAS P.

; REGISTRATION NUMBER: 38,548

; REFERENCE/DOCKET NUMBER: 9001-0035

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 325-7812

; TELEFAX: (650) 325-7823

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 49 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-019-010-4

alignment\_scores:

Quality: 290.00 Length: 49

Ratio: 5.918 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-12 x US-09-019-010-4 ..

Align seg 1/1 to: US-09-019-010-4 from: 1 to: 49

```

25 CAGCATTTGAGACTAGGCGCTGCGCCCTGGCAGCGTTCTCAAGATTGGAG 74
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
1 GlnHisTyrSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 17

```

```

75 CTACGGCCCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTAGGCGCTGC 124

```

```

|||||
17 rtyrglyleuargproglyserglnhistrpserlyrglyleua 34
125 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTCGCGGT 171
34 rproglyserglyserglnasptrpserlyrglyleuargprogly 49

```

seq\_name: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:US-09-249-447A-7

seq\_documentation\_block:  
; Sequence 7, Application US/09249447A

```

; GENERAL INFORMATION:
; APPLICANT: Manns, Jack G.
; TITLE OF INVENTION: PASSIVE IMMUNIZATION AS A TREATMENT FOR HORMONE
; TITLE OF INVENTION: DEPENDENT DISORDERS
; FILE REFERENCE: 9001-0045
; CURRENT APPLICATION NUMBER: US/09/249,447A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/075,637
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 7
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GnRH-2, Fig.
US-09-249-447A-7

```

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-12 x US-09-249-447A-7 ..

Align seg 1/1 to: US-09-249-447A-7 from: 1 to: 49

```

25 CAGCATTTGAGCTAGCGGCTCGGCTGAGCGGTTCTCAAGATTGAG 74
|||||
1 Glnhistrpserlyrglyleuargproglyserglyserglnasptrpse 17
75 CTACGGCTGCGGTCCGGTGGCTCTAGCCAGCATTTGAGCTAGCGCTGC 124
|||||
17 rtyrglyleuargproglyserglnhistrpserlyrglyleua 34
125 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTCGCGGT 171
34 rproglyserglyserglnasptrpserlyrglyleuargprogly 49

```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-305-924-11

seq\_documentation\_block:

```

; Sequence 11, Application US/09305924A
; GENERAL INFORMATION:
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; PRIOR FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 11
; LENGTH: 49
; TYPE: PRT

```

```

; ORGANISM: GnRH
US-09-305-924-11

```

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-12 x US-09-305-924-11 ..

Align seg 1/1 to: US-09-305-924-11 from: 1 to: 49

```

25 CAGCATTTGAGCTAGCGGCTCGGCTGAGCGGTTCTCAAGATTGAG 74
|||||
1 Glnhistrpserlyrglyleuargproglyserglyserglnasptrpse 17
75 CTACGGCTGCGGTCCGGTGGCTCTAGCCAGCATTTGAGCTAGCGCTGC 124
|||||
17 rtyrglyleuargproglyserglnhistrpserlyrglyleua 34
125 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTCGCGGT 171
34 rproglyserglyserglnasptrpserlyrglyleuargprogly 49

```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-306-689-11

seq\_documentation\_block:

```

; Sequence 11, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; TITLE OF INVENTION: ANIMALS
; FILE REFERENCE: 9001-0047
; CURRENT APPLICATION NUMBER: US/09/306,689B
; PRIOR FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/088,024
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 11
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
US-09-306-689-11

```

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-12 x US-09-306-689-11 ..

Align seg 1/1 to: US-09-306-689-11 from: 1 to: 49

```

25 CAGCATTTGAGCTAGCGGCTCGGCTGAGCGGTTCTCAAGATTGAG 74
|||||
1 Glnhistrpserlyrglyleuargproglyserglyserglnasptrpse 17
75 CTACGGCTGCGGTCCGGTGGCTCTAGCCAGCATTTGAGCTAGCGCTGC 124
|||||
17 rtyrglyleuargproglyserglnhistrpserlyrglyleua 34
125 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTCGCGGT 171
34 rproglyserglyserglnasptrpserlyrglyleuargprogly 49

```

```
seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:us-09-383-912-4
seq_documentation_block:
; Sequence 4, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016,22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-4

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-12 x US-09-383-912-4 ..

Align seg 1/1 to: US-09-383-912-4 from: 1 to: 49
25 CAGCATGGAAGCTACGGCTGCGCTGCGAGCGCTTCTCAAGATTGGAG 74
|||||
1 GlnIstIrpSerTyrGlyLeuAlaGyrProGlySerGlnIstIrpSer 17
|||||
75 CTACGGCTGCGCTGCGAGTGTACGACGATGGAAGCTACGGCTGC 124
|||||
17 rTyGlyLeuAlaGyrProGlySerGlnIstIrpSerTyrGlyLeuAla 34
|||||
125 GCGCTGCGAGCGGTACGACGATGGAAGCTACGGCTGCGCGGT 171
|||||
34 rGrProGlySerGlySerGlnIstIrpSerTyrGlyLeuAlaGyrProGly 49

seq_name: /cgn2_6/ptodata/2/paa/US086_COMB.pep:us-08-669-785-2
seq_documentation_block:
; Sequence 2, Application US/08669785
; GENERAL INFORMATION:
; APPLICANT: Betson, Fotini
; APPLICANT: Sebo, Peter
```

```
APPLICANT: Guiso, Nicole
TITLE OF INVENTION: Protective Epitopes Of Adenyli
TITLE OF INVENTION: Cyclase-haemolysin(AC-Hly), Their Application To
TITLE OF INVENTION: The Treatment Of To The Prevention Of Bordetella Infections
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farbow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,785
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356,0072-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1706 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-669-785-2

alignment_scores:
Quality: 212.00 Length: 512
Ratio: 0.763 Gaps: 21
Percent Similarity: 54.297 Percent Identity: 22.266

alignment_block:
US-09-306-689-12 x US-08-669-785-2 ..

Align seg 1/1 to: US-08-669-785-2 from: 1 to: 1706
265 AGCTACGCGCTGCGCTGCGAGCGGTAGCCAGATTGAGCTACGGCT 314
|||||
349 AlaTyrGlyVal...AlaGlyLysSerLeuPheAspArgGlyLeuGly 364
|||||
315 GCGTCCGGGTGGATCTAGCTTCCCAAAAGTGGGCAAAAATTTATCC 364
|||||
364 AlaLrProGlyValrProSerGlyArgSerLysrPheSerProAspValLeuG 381
|||||
365 TCTATATTCGCCAAATTTACCAATATGATAGTGAACAAGTAATGTTT 414
|||||
381 LurHValProAlaSer.....ProGlyLeu 389
|||||
415 CAGGATTATGTCAAAGCGCGGAGAGAGTTGGGATGAGTACAAGA 464
|||||
390 Arg.....ArgProSerLeuGly...AlaValGlyValrGly 400
|||||
465 AGAAGC.....ATAATATTGCAACAGCTCAACAGCTT 499
|||||
400 nAspSerGlyTyrAspSerLeuAspGlyValGlySerArgSerPheSerL 417
|||||
500 TAGCAGAGATTCAACCGCTATTGGCTTAACTGAGCGCATGTGTTA 549
|||||
417 euGlyGlyValSerAspMet.....AlaVal 426
|||||
550 TCCGCTCCACAAATTTGATTAATTCCTACAGAAAATAAGCGCCCAAGC 599
```



```

328 TCTAGCTCCCAAAAGTGGGCAAAAATATTCCTATATT...CC 374
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
60 MetSerLeuPro...ThrProAlaProArgValThrValTyrValAspPr 75
375 CCAAAATTAC 384
|||
75 OProAlaTyr 78

```

seq\_name: /cgn2\_6/prodata/2/paa/US60\_COMB.pep:US-60-120-454-23

```

seq_documentation_block:
; Sequence 23, Application US/60120454
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Yule, Terecila D
; APPLICANT: Martind, Serge A
; APPLICANT: Durtsehl, Becky A
; TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A
; FILE REFERENCE: PC10202
; CURRENT APPLICATION NUMBER: US/60/120,454
; EARLIER FILING DATE: 1999-02-17
; EARLIER APPLICATION NUMBER: N/A
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 23
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 4GRRH-tmgd
; OTHER INFORMATION: encoded by pOE-GRRH.gd.
US-60-120-454-23

```

alignment\_scores:

Quality:	208.50	Length:	70
Ratio:	4.088	Gaps:	5
Percent Similarity:	72.857	Percent Identity:	58.571

alignment\_block:

US-09-306-689-12 x US-60-120-454-23 ..

Align seg 1/1 to: US-60-120-454-23 from: 1 to: 411

```

178 CAGCATTTGGAGCTAGCGCTCGCCCTGGCGAGCGCTTCTCAAGATTGAG 227
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
19 GluHisTrpSerTyrGlyLeuArgProGly.....GluHisTrpSe 32
228 CTAGCGCGCTGGCGGTGGCTGCTAGCCAGCATTTGAGCTAGCGCGCTGC 277
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
32 TTYrGlyLeuArgProGly.....GluHisTrpSerTyrGlyLeuA 46
278 GCCCTGGCGGGTAGCCAGATTGGAGCTAGCGCTCGCTCGCGGTGGA 327
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
46 rProGly.....GluHisTrpSerTyrGlyLeuArgProGlySer 59
328 TCTAGCTCCCAAAAGTGGGCAAAAATATTCCTATATT...CC 374
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
60 MetSerLeuPro...ThrProAlaProArgValThrValTyrValAspPr 75
375 CCAAAATTAC 384
|||
75 OProAlaTyr 78

```

seq\_name: /cgn2\_6/prodata/2/paa/US086\_COMB.pep:US-08-669-785-4

```

seq_documentation_block:
; Sequence 4, Application US/08669785
; GENERAL INFORMATION:
; APPLICANT: Betsou, Fotini
; APPLICANT: Sebo, Peter

```

```

; APPLICANT: Gulso, Nicole
; TITLE OF INVENTION: Protective Epitopes Of Adeny1
; TITLE OF INVENTION: Cyclase-Hemolysin(AC-HLY). Their Application To
; TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella Infections
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,785
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356, 0072-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4400
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1705 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-669-785-4

```

alignment\_scores:

Quality:	206.50	Length:	372
Ratio:	1.059	Gaps:	11
Percent Similarity:	52.419	Percent Identity:	23.387

alignment\_block:

US-09-306-689-12 x US-08-669-785-4 ..

Align seg 1/1 to: US-08-669-785-4 from: 1 to: 1705

```

300 TTGAGAGCTAGCGCTCGCGGTGGATCTAGCTCCCAAAAGTGGG 349
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
380 LeuGluThrValProAlaSer.ProGlyLeuArgArgProSerLeuGly 396
350 CAAAAAATATTCCTATATTCGCCAAATATTCATATGATGATCGAA 399
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
396 laValGlu.....ArgGlnAspSerGlyTyrAspSerLeu 407
400 CAAGTATAGGTTTACAGATTAGTCAAAAGCGCGCAAGAGTGGGAT 449
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
408 AspGlyValGlySerArgSerPhe..... 415
450 TGAGGTCAAAAGAGAAAGCAATTAATTCGAACAGCTCAACAGATT 499
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
416 .....SerL 417
500 TAGCGACATTCAAAACCGCTATTGGCTTAAGTAGCGCTGATGTGTA 549
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
417 euGlyGluValSerAspMet.....AlaAlaVal 426
550 TCGCGTCACAAATGATTAATGCTAGACAGAAACTAAGCAGCGCAAGC 599
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
427 GluAlaAlaGluLeuGluMetThrArgGlnValLeuHisAlaGlyAla 443
600 ATTAGGTTCTGCCGAAGCATTGTACAAAAATGCAAATATAA.....GCCA 643

```



[illegible]

```

:
:
:
TITLE OR INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
:
: FILE REFERENCE: IGH-200440
:
: CURRENT APPLICATION NUMBER: PCT/US00/22121
:
: CURRENT FILING DATE: 2000-08-14
:
: PRIOR APPLICATION NUMBER: 09/026,276
:
: PRIOR FILING DATE: 1998-02-19
:
: PRIOR APPLICATION NUMBER: US 09/374,721
:
: PRIOR FILING DATE: 1999-08-13
:
: NUMBER OF SEQ ID NOS: 35
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 35
:
: LENGTH: 40
:
: TYPE: PRT
:
: ORGANISM: Artificial Sequence
:
: FEATURE:
:
: OTHER INFORMATION: Description of Artificial Sequence: polypeptide
PCT-US00-22121-35

alignment_scores:
Quality: 202.50 Length: 61
Ratio: 5.192 Gaps: 3
Percent Similarity: 63.934 Percent Identity: 63.934

alignment_block:
US-09-306-689-12 x PCT-US00-22121-35 ..

Align seg 1/1 to: PCT-US00-22121-35 from: 1 to: 40

25 CAGCATTTGAGCTACGGCCCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGly..... 10

75 CTACGGCCCTGGCGTCCGGGTGGCTCTGACGACGATTTGAGACTACGGCCCTGC 124
|||||
11 .....GlnHisTrpSerTyrGlyLeuArg 18

125 GCCCTGGCAGCGGCTAGCCAGATTGGAGCTACGCGCCTGCGCCGGGTGGA 174
|||||
18 rgrProGly.....GlnHisTrpSerTyrGlyLeuArgProGly... 30

175 TCTCAGCATTTGAGCTACGGCCCTGCGCCCTGGC 207
|||||
31 ...GlnHisTrpSerTyrGlyLeuArgProGly 40

seq_name: /cgr2_6/ptodata/2/paa/US090_COMP.pep:US-09-026-276-35

seq_documentation_block:
: Sequence 35, Application US/09026276
: GENERAL INFORMATION:
:
: APPLICANT: Kenten, John H
:
: APPLICANT: Tremontano, Alfonso
:
: APPLICANT: Pilon, Aprille L
:
: APPLICANT: Lohnas, Gerald F
:
: TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
:
: FILE REFERENCE: U.S. Patent Application No. 09/026,276
:
: CURRENT APPLICATION NUMBER: US/09/026,276
:
: CURRENT FILING DATE: 1998-02-19
:
: NUMBER OF SEQ ID NOS: 35
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 35
:
: LENGTH: 40
:
: TYPE: PRT
:
: ORGANISM: Porcine
US-09-026-276-35

alignment_scores:
Quality: 202.50 Length: 61
Ratio: 5.192 Gaps: 3
Percent Similarity: 63.934 Percent Identity: 63.934

```

## alignment\_block:

US-09-306-689-12 x US-09-026-276-35 ..

Align seg 1/1 to: US-09-026-276-35 from: 1 to: 40

```

25 CAGCATTTGAGACTAGGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
   |||||||||||||||||||||||||||||||||||||||||||||||||||
1  GlnHistrpserTyrglyLeuAArgProgly..... 10

75 CTAGCGGCTCGGCTCGGCTGAGCTAGCCAGCATTTGAGACTAGCGGCTGC 124
   |||||||||||||||||||||||||||||||||||||||||||||||||||
11 .....GlnHistrpserTyrglyLeuA 18

125 GCCCTGGCAGCGGTAGCCAGATTGAGACTAGGCGCTGCGGCTGGA 174
   ||||||| || |||||||||||||||||||||||||||||||
18 rGProgly.....GlnHistrpserTyrglyLeuAArgProgly... 30

175 TCTCAGCATTTGAGACTAGGCGCTGCGCCCTGGC 207
   |||||||||||||||||||||||||||||||||||||||||||||||||||
31 ...GlnHistrpserTyrglyLeuAArgProgly 40

```

seq\_name: /cgn2.6/ptodata/2/paa/US093\_COMB.pep:US-09-374-721A-35

## seq\_documentation\_block:

```

; Sequence 35, Application US/09374721A
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H.
; APPLICANT: Roberts, Steven
; APPLICANT: Lohas, Gerald
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: CIP OF IGN-9601
; CURRENT APPLICATION NUMBER: US/09/374,721A
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 09/026,276
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 35
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; OTHER INFORMATION: antigen
US-09-374-721A-35

```

## alignment\_scores:

Quality:	202.50	Length:	61
Ratio:	5.192	Gaps:	3
Percent Similarity:	63.934	Percent Identity:	63.934

## alignment\_block:

US-09-306-689-12 x US-09-374-721A-35 ..

Align seg 1/1 to: US-09-374-721A-35 from: 1 to: 40

```

25 CAGCATTTGAGACTAGGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
   |||||||||||||||||||||||||||||||||||||||||||||||||||
1  GlnHistrpserTyrglyLeuAArgProgly..... 10

75 CTAGCGGCTCGGCTCGGCTGAGCTAGCCAGCATTTGAGACTAGCGGCTGC 124
   |||||||||||||||||||||||||||||||||||||||||||||||||||
11 .....GlnHistrpserTyrglyLeuA 18

125 GCCCTGGCAGCGGTAGCCAGATTGAGACTAGGCGCTGCGGCTGGA 174
   ||||||| || |||||||||||||||||||||||||||||||
18 rGProgly.....GlnHistrpserTyrglyLeuAArgProgly... 30

175 TCTCAGCATTTGAGACTAGGCGCTGCGCCCTGGC 207
   |||||||||||||||||||||||||||||||||||||||||||||||||||
31 ...GlnHistrpserTyrglyLeuAArgProgly 40

```



XX WPI: 1982-44213E/22 (44213E).

DR Microencapsulated polypeptide sustained release compsn. - useful

XX for release of luteinizing hormones etc. for fertility control

PT

XX Disclousure; Page 7; 27pp; English.

PS The sequence in P20121 is the consensus sequence of a hormonally

XX active polypeptide. The sequences in P20403-6 are specific examples of

CC these polypeptides. These polypeptides are water soluble and are

CC analogues of luteinizing hormone/releasing hormone (LH-RH). These

CC peptides are used within the scope of the invention in hormonally

CC active microcapsule formulations comprising at least one of these

CC polypeptides and optionally a polymer hydrolysis modifying agent

CC intimately mixed with or coated by a biocompatible, biodegradable

CC polymer which when administered to a mammal, will release a daily

CC amount of polypeptide effective for maintaining an hormonally

CC related condition over a predetermined period of time. The

CC specific hormonally related condition is the control of fertility

CC and physiologically related effects.

XX

SQ Sequence 10 AA:

Query Match 89.7%; Score 52; DB 3; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.0011;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

DB 2 hwsyalrpg 10

RESULT 2

P40404 P40404 standard; peptide; 10 AA.

XX

AC P40404;

XX

DT 04-FEB-1992 (first entry)

XX

DE Sequence of peptide in a nasal composition having luteinizing

DE hormone releasing hormone (LHRH) agonist activity.

XX

KW Abortion; inducer; fertility control; growth promoter;

XX chemical sterlant; prostate cancer therapy; hormone.

XX

XX Key Location/Qualifiers

FT Modified-site 1

FT Modified-site 6 /label= pyroglu

FT Modified-site 10 /label= 3-(2-naphthyl)-D-Ala

FT Modified-site 10 /label= Gly-NH2

XX

PN EP111841-A.

XX

PD 27-JUN-1984.

XX

XX 08-DEC-1983; 83EP-0112369.

XX

PR 10-DEC-1982; 82US-0448548.

XX

XX (SYNT ) SYNTX (USA) INC.

XX

XX Anlk ST;

XX

XX WPI: 1984-159888/26.

XX

PT Nasal compsn. contg. nona: or deca:peptide - having LHRH

PT activity, with bile acid (salt) surfactant for enhanced

PT absorption

XX Claim 4; Page 24; 33pp; English.

PS The inventors claim a nasal compsn. having LHRH agonist activity,

XX together with a bile acid or its salt as surfactant, in aq. soln.

CC Preld. peptides are described in US4234571 and are useful for

CC fertility control, growth enhancement, treatment of prostatic cancer,

CC inducing abortion. The compsn. comprises 0.005-5 mg/ml of the

CC peptide, with 0.2-5% w/v of the surfactant, together with the buffer

CC soln. A typical compsn. contains 0.05-4 mg/ml of a peptide and 0.5-4% sodium glycocholate. The molar ratio of bile acid to peptide is

CC equal to or greater than 20:1. An additional surfactant may be

CC included.

XX

SQ Sequence 10 AA:

Query Match 89.7%; Score 52; DB 5; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.0011;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

DB 2 hwsyalrpg 10

RESULT 3

P71316 P71316 standard; peptide; 10 AA.

XX

AC P71316;

XX

DT 30-APR-1991 (first entry)

XX

DE Sequence of luteinizing hormone releasing hormone (LHRH) analogues,

DE having D-3 pyridyl-alanine at position 6.

XX

XX Prostate cancer therapy; agonist.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Misc-difference 1 /label= pyro-glu

FT Misc-difference 6

FT Misc-difference 9 /label= 3-(3-pyridyl)-D-Ala

FT Misc-difference 9 /label= Pro-NHCH3, Pro-NHC2H5, Pro-NHC3H7

FT Misc-difference 10 /label= "if Gly (10) absent"

FT Misc-difference 10 /label= Gly-NH2

FT Misc-difference 10 /note= "or absent"

XX

PN US4642332-A.

XX

XX 10-FEB-1987.

XX

XX 26-APR-1985; 85US-0727711.

XX

PR 26-APR-1985; 85US-0727711.

XX

XX (TEXA ) UNIV TEXAS SYSTEM.

XX

XX Folkers K, Jie-Cheng X, Bowers CY;

XX

XX WPI: 1987-056353/08.

XX

XX Luteinizing hormone releasing hormone analogues - of greater

XX potency, having D-3-pyridyl-alanine or 3-quinolyl-D-alanine at

XX position 6

XX

XX Claim 1; column 11; 7pp; English.

CC These peptides are synthetic analogues of LHRH which are termed  
 CC super agonists because of their superior potency. They may be used  
 CC to treat prostate cancer. The peptides are synthesised by solid  
 CC phase techniques using benzhydrylamine resin as solid support.

XX Sequence 10 AA;

Query Match 89.7%; Score 52; DB 8; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.0011;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
 |||||  
 Db 2 hwsyalrpg 10

RESULT 4  
 P71226  
 ID P71226 standard; peptide; 10 AA.

XX P71226;

XX 29-APR-1991 (first entry)

DE Luteinising hormone releasing hormone analogue.

KW Luteinising hormone releasing hormone; (D-Ala6)-LHRH; cancer;  
 KW cytotoxic effects; bone marrow; agonist; radiotherapy; chemotherapy;

XX CH659392-A.

XX 30-JAN-1987.

XX 07-MAY-1984; 84CH-0002226.

XX 07-MAY-1984; 84CH-0002226.

XX (DEBI-) DEBIOPHARM SA.

XX Mathe G, Misset JL, Lemaigre G, Mauvernay RY, Schally AV;

DR WPI; 1987-037621/06.

XX Compsns. inhibiting cytotoxic effects in cancer therapy - contain  
 PT luteinising and follicle stimulating hormone release hormone or  
 PT its analogues

PS Claim 2; page 2; 4pp; English.

CC The analogue is an agonist. It inhibits the cytotoxic effects of  
 CC cancer therapy on bone marrow cells, and protects the cells from the  
 CC undesirable effects of radio- and chemotherapy used to treat hormone-  
 CC dependent and -independent cancers. The compsns. may be in sustained  
 CC release form and may be given parenterally, esp. i.m.

XX Sequence 10 AA;

Query Match 89.7%; Score 52; DB 8; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.0011;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
 |||||  
 Db 2 hwsyalrpg 10

RESULT 5  
 P71229  
 ID P71229 standard; protein; 10 AA.

XX P71229;

XX 29-APR-1991 (first entry)

DE Luteinising hormone releasing hormone analogue.

KW Luteinising hormone releasing hormone; (D-Ser(Bu+6,des Gly-NH2 10)-LHRH;  
 KW cytotoxic effects; bone marrow; agonist; radiotherapy;

XX chemotherapy; cancer.

XX CH659392-A.

XX 30-JAN-1987.

XX 07-MAY-1984; 84CH-0002226.

XX 07-MAY-1984; 84CH-0002226.

XX (DEBI-) DEBIOPHARM SA.

XX Mathe G, Misset JL, Lemaigre G, Mauvernay RY, Schally AV;

XX WPI; 1987-037621/06.

XX Compsns. inhibiting cytotoxic effects in cancer therapy - contain  
 PT luteinising and follicle stimulating hormone release hormone or  
 PT its analogues

PS Claim 2; page 2; 4pp; English.

CC The analogue is an agonist. It inhibits the cytotoxic effects of  
 CC cancer therapy on bone marrow cells, and protects the cells from the  
 CC undesirable effects of radio- and chemotherapy used to treat hormone-  
 CC dependent and -independent cancers. The compsns. may be in sustained  
 CC release form and may be given parenterally, esp. i.m.

XX Sequence 10 AA;

Query Match 89.7%; Score 52; DB 8; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.0011;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
 |||||  
 Db 2 hwsyalrpg 10

RESULT 6

XX P71363 standard; peptide; 10 AA.

XX P71363;

XX 26-MAY-1991 (first entry)

DE Sequence of a luteinising hormone-releasing hormone (LHRH) analogue  
 DE used in a compsn. to treat infertility.

XX Infertility treatment; impotency; spermatogenesis; androgen;  
 KW steriliant abortion.

XX Key Location/Qualifiers

FT Misc-difference 1 /label= (pyro)Glu

FT Misc-difference 5 /label= Tyr-3-(naphthyl)

FT Misc-difference 6 /label= D-Ala

FT Misc-difference 10 /label= Gly-NH2

XX USA675189-A.

PD 23-JUN-1987.  
 XX  
 CC 08-FEB-1985; 85US-0699715.  
 XX  
 PR 11-DEC-1981; 81US-0329832.  
 PR 18-NOV-1980; 80US-0207864.  
 PR 08-FEB-1985; 85US-0699715.  
 XX  
 PA (SYNTH ) SYNTEX (USA) INC.  
 XX  
 PI Kent JS, Lewis DH, Sanders LM, Tice TR;  
 DR WPI; 1987-192071/27.  
 XX  
 PT Sustained release compsn. microcapsules - contain luteinising  
 PT hormone releasing hormone analogues and  
 PT (poly-lactide-co-glycolide) polymer, to treat infertility  
 XX  
 PS Claim 9; column 18; 11pp; English.  
 XX  
 CC Low doses of the compsn. are used to treat infertility in females  
 CC and impotency and to stimulate spermatogenesis and androgen prodn.  
 CC in males. High doses block ovulation and suppress spermatogenesis;  
 CC and may be used to promote wt. gain in animals, to stimulate  
 CC abortion in animals and as a chemical sterilant. (See US4234571).  
 CC For agonist fertility control, the microcapsules release the drug to  
 CC dose 0.01-100 (pref. 0.1-5.0) micrograms/kg/day.  
 CC  
 SQ Sequence 10 AA;

Query Match 89.7%; Score 52; DB 8; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.0011;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 DB 2 hwsyalrpg 10

RESULT 7  
 P83021  
 ID P83021 standard; protein; 10 AA.  
 XX  
 AC P83021;  
 XX  
 DT 11-DEC-1990 (first entry)  
 XX  
 DE Example of gonadoliberin peptide analogue.  
 XX  
 KM Gonadoliberin; agonists; antagonists; fertility; amenorrhoea;  
 KM oligospermia; contraceptives; cancers.  
 OS  
 OS synthetic.  
 XX  
 PN EP263521-A.  
 XX  
 PD 13-APR-1988.  
 XX  
 PF 08-OCT-1987; 87EP-0114702.  
 XX  
 PR 09-OCT-1986; 86DE-3634435.  
 XX  
 PA (FARRH ) HOECHST AG.  
 PI Konig W, Sandow JK, Kolar C;  
 DR WPI; 1988-099948/15.  
 XX  
 PT New peptide derivs. contg. glycosyl gp. - useful as gonadoliberin  
 PT agonists or antagonists  
 PS Claim 1; Page 40; 43pp; German.

XX This peptide is an agonist or antagonist of gonadoliberin. It, and  
 CC its salts, is useful as an agonist for promoting fertility e.g. for  
 CC treating amenorrhoea, luteal insufficiency, oligospermia and retarded  
 CC puberty. It is also useful as a contraceptive and for treating  
 CC steroid-dependent cancers. As an antagonist it is useful for  
 CC treating gonadotropin- and steroid-dependent diseases.  
 CC The N-terminal has an H gp. or a 1-7C acyl attached. Residues 1, 2  
 CC and 3 are opt. ring subst. by one or two of Br, Cl, F, NO<sub>2</sub>, NH<sub>2</sub>, Me  
 CC or OMe. Residues 6 and 7 opt. have a partially protected glycosyl  
 CC attached which increases their water-solubility without impairing  
 CC their activity.  
 CC See also P82831 and P83017-20.  
 XX  
 SQ Sequence 10 AA;

Query Match 89.7%; Score 52; DB 9; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.0011;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||||  
 DB 2 hwsyalrpg 10

RESULT 8  
 R10699  
 ID R10699 standard; peptide; 10 AA.  
 XX  
 AC R10699;  
 XX  
 DT 22-APR-1991 (first entry)  
 XX  
 DE LHRH analogue 3.  
 XX  
 KM Luteinising hormone releasing hormone.  
 XX  
 OS Synthetic.  
 OS  
 FT Key Location/Qualifiers  
 FT Modified-site 1 /label= pyro-Glu  
 FT Modified-site 4 /label= N-Me-Ser  
 FT Modified-site 6 /label= 3-(2-naphthyl)-D-Ala  
 FT  
 PN EP413209-A.  
 XX  
 PD 20-FEB-1991.  
 XX  
 PF 02-AUG-1990; 90EP-0114829.  
 XX  
 PR 10-JUL-1990; 90US-0548512.  
 PR 07-AUG-1989; 89US-0390572.  
 XX  
 PA (ABBO ) ABBOTT LABORATORIES.  
 XX  
 PI Oxford R, Greer J;  
 DR WPI; 1991-052535/08.  
 XX  
 PT New pseudo nona- and deca-peptide derivs. of LHRH - used to  
 PT modulate level of sex hormones in males and females  
 XX  
 PS Claim 4; Page 74; 78pp; English.

CC The peptide is one of several examples of compounds derived from the  
 CC sequence of LHRH in which the nitrogen atom of at least one of the  
 CC amide bonds has been alkylated. The peptides are used to modulate  
 CC levels of sex hormones in males and females. They are stable  
 CC against proteolytic enzymes and biologically potent after oral admin.

CC in animals and humans. See also R10697-R10698.  
 XX Sequence 10 AA:

Query Match 89.7%; Score 52; DB 12; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.0011;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||| ||||  
 Db 2 hwsyalrpg 10

## RESULT 9

R15714 ID R15714 standard; Protein; 10 AA.

AC R15714;

DT 24-JAN-1992 (first entry)

DE Peptide #2 with homology to LHRH.

KW Luliberin.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1

FT Modified-site 6

FT Modified-site 9

FT Modified-site 10

FT Modified-site 10

PN W09116343-A.

PD 31-OCT-1991.

PF 22-APR-1991; 91WO-FR00332.

PR 23-APR-1990; 90FR-0005147.

PA (INRM ) INSERM INST NAT SANTE.

PI Gautron J, Patton E, Kordon C, Bauer K;

DR WPI; 1991-339753/46.

PT New peptide homologous with luteinising hormone-releasing hormone

PT - used to treat gynaecological conditions, cancer of gonads and

PT sec. sexual organs, psychiatric conditions and in assays

PS Claim 6; Page 50; 83pp; French.

CC The C-terminal residue (Gly-CO-NH2) can be replaced by ethylamide.

CC This peptide and fragments of it (i.e. amino acids 4-10, 5-10, 6-10

CC and 7-10) are agonists and antagonists of LHRH. They are useful for

CC treating e.g. precocious or delayed puberty, psychiatric disorders

CC esp. those of the libido or sexual aggression, etc. In addition they

CC are useful for functional exploration of the hypothalamus-hypophyseal

CC axis and for radioimmunological or biological assay (of LH, FSH and

CC steroid levels) in biological fluids and biopsy samples.

Query Match 89.7%; Score 52; DB 12; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.0011;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
 |||| ||||  
 Db 2 hwsyalrpg 10

## RESULT 10

R15715 ID R15715 standard; Protein; 10 AA.

AC R15715;

DT 24-JAN-1992 (first entry)

DE Peptide #3 with homology to LHRH.

KW Luliberin.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1

FT Modified-site 6

FT Modified-site 9

FT Modified-site 10

FT Modified-site 10

PN W09116343-A.

PD 31-OCT-1991.

PF 22-APR-1991; 91WO-FR00332.

PR 23-APR-1990; 90FR-0005147.

PA (INRM ) INSERM INST NAT SANTE.

PI Gautron J, Patton E, Kordon C, Bauer K;

DR WPI; 1991-339753/46.

PT New peptide homologous with luteinising hormone-releasing hormone

PT - used to treat gynaecological conditions, cancer of gonads and

PT sec. sexual organs, psychiatric conditions and in assays

PS Claim 6; Page 50; 83pp; French.

CC The C-terminal residue (Gly-CO-NH2) can be replaced by ethylamide.

CC This peptide and fragments of it (i.e. amino acids 4-10, 5-10, 6-10

CC and 7-10) may be agonists or antagonists of LHRH. They are useful for

CC treating e.g. precocious or delayed puberty, psychiatric disorders

CC esp. those of the libido or sexual aggression, etc. In addition they

CC are useful for functional exploration of the hypothalamus-hypophyseal

CC axis and for radioimmunological or biological assay (of LH, FSH and

CC steroid levels) in biological fluids and biopsy samples.

OY 2 HWSYXLRPG 10

Db 2 hwsyslrpg 10

## RESULT 11

R29612 ID R29612 standard; peptide: 10 AA.

AC R29612;

DT 23-APR-1993 (first entry)

DE Goserelin, contg. C-terminal aza-amino acid amide.

KW LHRH; luteinising hormone releasing hormone; analogue; prostate; cancer; breast; gynaecological conditions.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1 /note= "pyroglu"

FT Misc-difference 6 /note= "D-Ser(But)"

FT Modified-site 10 /note= "amidated azagly"

EP518656-A.

PD 16-DEC-1992.

PF 11-JUN-1992; 92EP-0305341.

PR 14-JUN-1991; 91GB-0012859.

PA (ICIL ) IMPERIAL CHEM IND PLC.

XX Hayward CF;

DR WPI; 1992-417511/51.

PT Solid phase synthesis of peptide(s) contg. aza-aminoacid - partic. goserelin, by assembling aminoacid(s) of peptide except C terminal aza aminoacid and cleaving peptide from support with hydrazine, etc.

XX PS Disclosure; Page 2; 4pp; English.

CC Goserelin contg. a C-terminal aza-amino acid amide was prepd. by a novel method comprising assembling all the amino acids of the peptide except the C-terminal aza-amino acid by conventional solid CC phase synthesis, cleaving the peptide from the support with CC hydrazine and reacting the released hydrazine with a cyanate ion. CC The peptide also contains a tert-butyloxy gp. in its side chain. CC The peptide is an LHRH analogue useful in the treatment of prostate cancer, breast cancer and certain gynaecological conditions.

XX SQ Sequence 10 AA;

Query Match 89.7%; Score 52; DB 13; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.0011; Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

Db 2 hwsyslrpg 10

## RESULT 12

R29703

ID R29703 standard; peptide: 10 AA.

AC R29703;

DT 23-APR-1993 (first entry)

DE Goserelin contg. aza-amino acid amide.

KW Ester; LHRH; luteinising hormone releasing hormone; prostate; cancer; breast; gynaecological conditions.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1 /label= pglu

FT Misc-difference 6 /note= "D-Ser(But)"

FT Modified-site 10 /note= "amidated azagly"

EP518655-A.

PD 16-DEC-1992.

PF 11-JUN-1992; 92EP-0305340.

PR 14-JUN-1991; 91GB-0012825.

PA (ICIL ) IMPERIAL CHEM IND PLC.

XX Cotton R, Giles MB;

DR WPI; 1992-417510/51.

PT New N-protected-aza-amino-acyl ester(s) - are used in solid phase synthesis of peptide(s) contg. C-terminal aza-aminoacid, partic. goserelin

XX PS Disclosure; Page 2; 7pp; English.

CC Goserelin contg. an aza amino acid was prepd. by a novel solid phase synthesis method comprising reacting an active ester or imidazolone of an N-protected aza-amino acid with an appropriate reactive solid CC support, carrying out further conventional solid phase peptide CC synthesis steps to add sequentially further amino acids, to form a CC peptide with the required amino acid sequence bound to a solid CC support, cleaving the peptide from its support and opt. reacting the CC prod. with hydrazine to remove any acyl gps. formed during synthesis. CC The peptide is a synthetic analogue of LHRH and is useful in the CC treatment of prostate cancer, breast cancer and certain gynaecological conditions.

XX SQ Sequence 10 AA;

Query Match 89.7%; Score 52; DB 13; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.0011; Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

Db 2 hwsyslrpg 10

## RESULT 13

W22391

ID W22391 standard; peptide: 10 AA.

AC W22391;

DT 20-JAN-1998 (first entry)

DE Gonadotropin releasing hormone [(3-Mercaptopropionyl)-Gln1,dala6]GnRH.



KW	Gonadotropin releasing hormone; vaccine; immuno-sterilisation;
RW	fertility control; steroid hormone stimulated tumour; endometriosis;
KX	GnRH.
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1 /note= "Attached to HSC <sub>2</sub> H <sub>2</sub> C(ONH)-"
FT	Misc-difference 6 /note= "D-form residue"
FT	Modified-site 10 /note= "In amide form"
XX	
PN	W097J5316-A1.
XX	
PD	01-MAY-1997.
XX	
PE	23-OCT-1996; 96WO-US16950.
XX	
PR	13-FEB-1996; 96GB-0002951.
PR	27-OCT-1995; 95GS-0005905.
XX	
PA	(MERI ) MERCK & CO INC.
XX	
PI	Lombardo VK, Marburg S, Tolman RL;
XX	
DR	WPJ; 1997-258759/23.
PT	New gonadotropin releasing hormone conjugates - used as vaccines for
PT	immuno-sterilisation of animals, fertility control or for treating
PT	steroid hormone stimulated tumours or conditions.
XX	
PS	Disclosure; Page 33; 75pp; English.
XX	
CC	A gonadotropin releasing hormone conjugate has been developed which has
CC	the formula: Z-[L1-Rm-(Y1(L2-X)-An)q-Y2(L2-X)-Ap.(OH)r, where A = an
CC	amino acid selected from Gly, Ser, Thr, beta-Ala, and Ala, with the
CC	proviso that at least one A is Ser or Thr; L1 = a linker optionally
CC	attached to an internal marker; L2 = a linker; X = a gonadotropin
CC	releasing hormone (GnRH) having a free sulphydryl group; Y1, Y2 = Lys
CC	or ornithine (Orn); Z = an immunogenic carrier protein; m = 0-3; n = 5
CC	or 6; p = 0-1; q = 1 or 2; r = 1-3. They are useful as vaccines
CC	for use in immuno-sterilising animals, fertility control, and for
CC	treatment of steroid hormone stimulated tumour or conditions such as
CC	endometriosis. The conjugates employ a hydrophilic linear scaffold as a
CC	means to increase loading of desirable hapten onto a carrier protein.
CC	Compared to conventional GnRH conjugates, these provide more GnRH per
CC	carrier molecule while utilising a minimal number of carrier molecule
CC	attachment sites. The present sequence represents [(3-Mercaptopropionyl)
CC	-Gln1,dAla6]GnRH which is used in an example of the present invention.
CC	The novel vaccine compositions provide a high titre of anti-GnRH
CC	antibody and produce minimal injection site lesions.
XX	
SO	Sequence 10 AA:
OY	Query Match 89.7%; Score 52; DB 18; Length 10;
	Best Local Similarity 88.9%; Pred. No. 0.0011;
Matches	8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB	2 HMSXYLRPG 10       2 hwsyalrpg 10
RESULT 14	
ID	W16391 standard; peptide; 10 AA.
XX	
AC	W16391;
XX	
DT	14-JAN-1998 (first entry)

XX	Gonadotropin releasing hormone analogue; preparation 4.
DE	
XX	
KW	Gonadotropin releasing hormone; immunogen; Pseudomonas exotoxin;
KM	antibody; fertility; aggressive behaviour; tumour; steroid hormone;
KW	endometriosis; vaccine; sterility; breast cancer; prostatic cancer.
XX	
OS	Synthetic.
xx	
FH	Key
FT	Modified-site
FT	1 /note= "Attached to the linker HSCH2CH2CONH-"
FT	Misc-difference
FT	6 /note= "D-form residue"
FT	Modified-site
FT	10 /note= "In amide form"
xx	
PM	WO9715325-A1.
xx	
PD	01-MAY-1997.
xx	
PF	23-OCT-1996; 96WO-US17008.
xx	
PR	13-FEB-1996; 96GB-0002878.
PR	27-OCT-1995; 95US-0008018.
xx	
PA	(MERI ) MERCK & CO INC.
xx	
PI	Hickey GJ, Mohn RL;
xx	
DR	WPI: 1997-258766/23.
xx	
PT	Immunogen comprising gonadotropin releasing hormone and Pseudomonas
PT	exotoxin - generates antibodies that control fertility, reduce
PT	aggressive behaviour, and treat tumours responsive to steroid
PT	hormones or endometriosis
xx	
PS	Disclosure; Page 34; 63pp; English.
CC	A method has been developed for the generation of anti-gonadotropin
CC	releasing hormone (GnRH) antibodies (Ab) by administration of an
CC	immunogen consisting of GnRH associated with pseudomonas exotoxin (PE)
CC	or its variants. The present sequence represents a GnRH analogue,
CC	preparation 4, used in the preparation of GnRH vaccines
CC	containing this immunogen are used to control fertility; reduce
CC	undesirable behaviour (e.g. aggression) induced by reproductive
CC	hormones and to treat (a) tumours that respond to steroid sex hormones
CC	(e.g. breast and prostatic cancers) or (b) endometriosis. Particularly
CC	the immunogen is used to sterilise animals, e.g. to improve feed
CC	efficiency and carcass quality, and to eliminate boar taint. PG is an
CC	effective immunogenic carrier for GnRH, providing high titres of Ab.
xx	
SO	Sequence 10 AA:
OY	2 HWSYXLRPG 10       DB 2 hwsyalrpg 10
RESULT 15	
W94891	Query Match 89.7%; Score 52; DB 18; Length 10; Best Local Similarity 88.9%; Pred. No. 0.0011;
ID. W94891 standard; peptide; 9 AA.	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps
AC W94891;	
AC	
DT 11-MAY-1999 (first entry)	
XX LHRH peptide fragment.	

```

XX LHRH; immune response; luteinising hormone releasing hormone; DT;
KM diphtheria toxoid; castrating; oestrus cycling; aggression; breast;
KM sexual activity; organoleptic; livestock; cell growth; malignant;
KM prostate; ovarian; oncofoetal; hyperplastic; pregnancy;
KM endometriosis; inflammatory response.
XX
XX Homo sapiens.
XX
XX MO9902180-A1.
XX
XX 21-JAN-1999.
XX
XX 09-JUL-1998; 98MO-AU00532.
XX
XX 09-JUL-1997; 97AU-0007768.
XX
XX (CSLC-) CSL LTD.
XX
XX McNamara MK;
XX
XX WPI; 1999-120511/10.
XX
XX New immunogenic leutenising hormone releasing hormone compositions -
XX comprise LHRH conjugated to diphtheria toxoid and adsorbed to an
XX ionic polysaccharide, used to inhibit reproductive function in
XX animals
XX
XX Example 3; Page 30; 41pp; English.
XX
XX The invention relates immunogenic composition for eliciting an immune
XX response to luteinising hormone releasing hormone (LHRH). The composition
XX comprises a LHRH-diphtheria toxoid (DT) conjugate adsorbed to an ionic
XX polysaccharide. The LHRH-DT compositions can be used for eliciting an
XX immune response to LHRH, for castrating an animal, for regulating an
XX oestrus cycling in a female animal or for inhibiting characteristics
XX induced by the sexual maturation of an animal, e.g. aggression or sexual
XX activity. They can also be used for achieving production gains in
XX livestock, e.g. reduction or elimination of unwanted organoleptic
XX characteristics from the meat of livestock. They can also be used for
XX inhibiting the growth of cells which are regulated directly or indirectly
XX by LHRH, e.g. malignant breast cells, malignant prostate cells, malignant
XX ovarian cells, malignant oncofoetal cells or hyperplastic cells. They can
XX also be used for down-regulating the libido of an animal. They can also
XX be used for inhibiting pregnancy, prostate enlargement, endometriosis or
XX inflammatory responses. The LHRH compositions induce a more effective
XX immune response against LHRH than the LHRH-carrier adjuvant compositions.
XX The effective immune response against LHRH results in prevention of the
XX release of the hormones LH and FSH from the anterior pituitary. Sequences
XX W94890-93 are peptide derivatives of LHRH.
XX
XX Sequence 9 AA:
SO

```

```

Query Match      87.9%; Score 51; DB 20; Length 9;
Best Local Similarity 88.9%; Pred. No. 2.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 HWSYXLRPG 10
   ||||| |||||
Db 1 hwsyglrpg 9

```

```

Search completed: March 2, 2001, 10:53:45
Job time: 910 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:47 ; Search time 66.16 Seconds  
(without alignments)  
359.200 Million cell updates/sec

Title: US-09-306-689-13

Perfect score: 3591  
Sequence: 1 MATVIDRSDHWSYGLRPGSG.....GLRPGSGSDHWSYGLRPGGS 695

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_36: \*  
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT: \*  
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT: \*  
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT: \*  
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT: \*  
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT: \*  
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT: \*  
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT: \*  
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT: \*  
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT: \*  
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT: \*  
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT: \*  
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT: \*  
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT: \*  
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT: \*  
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT: \*  
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT: \*  
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT: \*  
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT: \*  
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT: \*  
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT: \*  
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3591	100.0	695	19	W79573
2	3591	100.0	695	21	Y58361
3	3591	100.0	695	21	Y58133
4	2642.5	73.6	544	19	W79570
5	2633.5	73.3	544	17	W03943
6	2426	67.6	977	17	W03942
7	2426	67.6	977	19	W79569
8	2354	65.6	490	20	Y33929
9	2173	60.5	936	14	R34547
10	2131	59.3	1069	15	R52748
11	2131	59.3	1069	18	W13867
12	2126	59.2	926	12	R14482

13	2126	59.2	926	14	R34545
14	2126	59.2	926	15	R50291
15	2126	59.2	926	17	W03945
16	2126	59.2	926	19	W79568
17	2117.5	59.0	1098	13	R22103
18	2117.5	59.0	1098	15	R52747
19	2117.5	59.0	1098	18	W13866
20	2116	58.9	924	14	R42385
21	2116	58.9	924	14	R42380
22	2116	58.9	924	14	R42378
23	2112	58.8	943	14	R34546
24	2110	58.8	951	14	R34548
25	2108.5	58.7	953	11	R07167
26	2108.5	58.7	953	14	R43885
27	2108.5	58.7	953	15	R60072
28	2105.5	58.6	953	12	R15159
29	2103	58.6	924	12	R10889
30	1650.5	46.0	1334	15	R50280
31	1650.5	46.0	1403	19	W79574
32	1644.5	45.8	1334	12	R14481
33	1551.5	43.2	1403	12	R10890
34	1524	42.4	934	17	W07637
35	1264.5	35.2	956	12	R12561
36	1264.5	35.2	956	18	W22156
37	1264.5	35.2	956	21	Y51410
38	938.5	26.1	1049	18	W22159
39	938.5	26.1	1049	21	Y51412
40	938.5	26.1	1244	15	R54781
41	848.5	23.6	1022	18	W22152
42	848.5	23.6	1022	21	Y51406
43	828.5	23.1	1023	16	R76991
44	403	11.2	758	17	R86998
45	380.5	10.6	323	12	R11187

#### ALIGNMENTS

RESULT 1	
ID W79573	W79573 standard; Protein; 695 AA.
XX AC W79573;	
DT 24-DEC-1998 (first entry)	
XX DE LKT-GNRH chimeric protein.	
XX KW Chimeric: PCB122; LKT 111; GNRH: Gonadotropin releasing hormone; multimer;	
KW cytotoxic activity; antigen presentation; Immune response; vaccine;	
KW tumour.	
XX OS Synthetic.	
XX PN W09806848-A1.	
XX PD 19-FEB-1998.	
XX PF 08-AUG-1997;	97WO-CA00559.
XX PR 09-AUG-1996;	96US-0694865.
XX PI (UYSA-) UNIV SASKATCHEWAN.	
XX PI Manns JG, Potter AA;	
XX DR WPI: 1998-159540/14.	
XX DR N-PSDB: V61535	
XX DX Leukotoxin carrier	
XX PT Chimeric protein of leukotoxin and gonadotropin releasing hormone	
PT useful for, e.g. preparation of vaccines for reduction of incidence	
PT of mammary tumours in mammals	
XX	

PS- Claim 9; Figure 9.1-6; 118pp; English.

XX The present sequence represents the LKT-GnRH chimeric protein from  
 CC PC8122. This plasmid contains the LKT 111 polypeptide fused to sixteen  
 CC copies of the GnRH peptide, in the pattern of: 8 copies of GnRH-LKT 111-8  
 CC copies of GnRH. This chimera lacks cytotoxic activity which enables  
 CC it to be an increase in antigen presentation and thus an optimal  
 CC immune response. The removal of this region also enables the truncated  
 CC LKT to be expressed at much higher levels and allows the amount of  
 CC antigen administered to be reduced. This chimeric protein comprises a  
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
 CC chimeric protein can be used as a vaccine to help reduce the incidence of  
 CC mammary tumours in a mammalian individual.

XX  
 XX Sequence 695 AA:

Query Match 100.0%; Score 3591; DB 19; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-233;  
 Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATVIDRSQHWSYGLRPGSGSDMSYGLRPGSSQHSYGLRPGSSQSDMSYGLRPGSSQ 60  
 DB 1 macvlrdsqhwysyglirpssgsqdwysyglirpssgsqdwysyglirpssgsq 60  
 QY HWSYGLRPGSGSDMSYGLRPGSSQHSYGLRPGSSQSDMSYGLRPGSSQSFPKTGAKKI 120  
 DB hwsyglirpssgsqdwysyglirpssgsqdwysyglirpssgsqdwysyglirpssgsq 120  
 QY 121 ILYIPONTYDTFQNGNGLDLYKAAEELGIEVOREERNNTATQTSIGTQTQATGTERG 180  
 DB 121 Ilyiponydytfeqngnlgldlykaaeelgievoreernnatqtslygtqtaigtterg 180  
 QY 181 IYVSAPQIDKLOKTKAGALGSFESIYONANKAKTVLSGSIISLVSYLAGMIDEALON 240  
 DB 181 IYVSAPQIDKLOKTKAGALGSFESIYONANKAKTVLSGSIISLVSYLAGMIDEALON 240  
 QY 241 NSNQHALLARAGLELNTSLIENIANSVKTLEDFEQOISQFSGSKLQNTKGLTGLDKLNIG 300  
 DB 241 nsnqhalakagletnslileniansvktldefeqisqfsgsklnkqlygtldklnig 300  
 QY 301 GUDKAGLGLDYSIGLSGATALVLDKKNASTKKGAGFELANOVYVGNITTKAVSYILA 360  
 DB 301 gldkaglgldysiglsataalvldkknastkkgagfelanqvynitlkavssyila 360  
 QY 361 ORVAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKKESYAERPKRTGYGDN 420  
 DB 361 orvaaglsstgpaalaaistvslaisplafagiadkfnhakslesyaerfkrtgygdnl 420  
 QY 421 LAEYQGTCTIDASVTAINMTALAAGVSAAMAADLTFEKVKNHLYTNSKKEKVTIQMW 480  
 DB 421 laeyggtctidasvtainmtalaagvsaamaadltfekvknhlytnskkekvtiqmw 480  
 QY 481 FRADRAKAYPNKATKDEKIEIITQNGERTISKOVDLIANGNGKITODELSKYVDNY 540  
 DB 481 freatdrakepnykatkdeieeiitqngeritlskqvddliakngkitldeelskyvddy 540  
 QY 541 ELTKHKNVTNSDLKLISSYSAFTSSNDSRNVLVAPTSMLOSLSSLOPARSGOHSYGL 600  
 DB 541 elltkhknvtnsdlklissysaftssndsrnvlvaptsmldqslsslofargsohsygl 600  
 QY 601 RPSGSGSDMSYGLRPGSSQHSYGLRPGSSQSDMSYGLRPGSSQHSYGLRPGSSQDW 660  
 DB 601 rpgsgsgdmsyglirpssgsqhwysyglirpssgsqdwysyglirpssgsqdw 660  
 QY 661 STGLRPGSSQHSYGLRPGSSQSDMSYGLRPGSS 695  
 DB 661 syglirpssgsqhwysyglirpssgsqdwysyglirpssgs 695

RESULT 2  
 Y58361  
 ID Y58361 standard; Protein; 695 AA.

XX  
 AC Y58361;  
 XX  
 DT 27-MAR-2000 (first entry)  
 XX  
 DE Leukotoxin/gonadotropin-releasing hormone fusion protein.  
 XX  
 KW Leukotoxin; gonadotropin-releasing hormone; GnRH; immunosterilisation;  
 KW immunosuppression; vaccine; feline; canine; equine; cervine.  
 XX  
 OS Chimeric - Pasteurella haemolytica.  
 OS Chimeric - Mammalia.  
 XX  
 PN WO962545-A2.  
 XX  
 PD 09-DEC-1999.  
 XX  
 XX 28-MAY-1999; 99WO-CA00493.  
 PF 04-JUN-1998; 98US-0088024.  
 PR 06-MAY-1999; 99US-0306689.  
 XX  
 PA (BIOS-) BIOSTAR INC.  
 XX  
 PI Robbins, SC;  
 XX  
 DR WPI: 2000-086857/07.  
 DR N-PSDB; Z55700.  
 XX  
 PT Hormone immunogens, analogues or antibodies used to manufacture  
 PT vaccines for suppression of reproductive behavior and fertility in  
 PT vertebrates -  
 XX  
 PS Claim 20; Fig 6A-6F; 88pp; English.  
 XX  
 CC This sequence represents a fusion protein comprising gonadotropin-  
 CC releasing hormone (GnRH) immunogens and a pasteurella  
 CC haemolytica leukotoxin (LKT) protein. The fusion protein comprises, in  
 CC the N to C-terminal direction, a synthetic peptide sequence (Y58364), an  
 CC eight copy GnRH multimer (composed of two copies of the 4xGnRH multimer  
 CC sequence of Y58363), the LKT protein (which functions as a carrier  
 CC protein), and a second eight copy GnRH multimer. The fusion protein may  
 CC be used in a vaccine composition for prepubertal administration to a  
 CC vertebrate subject to result in prolonged suppression of reproductive  
 CC behavior and/or fertility. GnRH immunogens, analogues or antibodies  
 CC that cross-react with endogenous GnRH of a vertebrate subject are used  
 CC to manufacture a composition or vaccine for immunosterilisation or  
 CC immunosuppression of feline, canine, equine or cervine subjects.  
 CC The vaccines are used to suppress reproductive behaviour and/or  
 CC fertility for at least 10 months. The prepubertal administration  
 CC results in a prolonged, long-term suppression of testicular development  
 CC and/or function in males, or a prolonged, long-term suppression of  
 CC ovarian development and/or function in females. The methods provide a  
 CC viable and desirable alternative to surgical forms of sterilisation that  
 CC are currently used.  
 XX  
 XX Sequence 695 AA:  
 XX  
 QY 1 MATVIDRSQHWSYGLRPGSGSDMSYGLRPGSSQHSYGLRPGSSQSDMSYGLRPGSSQ 60  
 DB 1 macvlrdsqhwysyglirpssgsqdwysyglirpssgsqdwysyglirpssgsq 60  
 QY HWSYGLRPGSGSDMSYGLRPGSSQHSYGLRPGSSQSDMSYGLRPGSSQSFPKTGAKKI 120  
 DB hwsyglirpssgsqdwysyglirpssgsqdwysyglirpssgsqdwysyglirpssgsq 120  
 QY 121 ILYIPONTYDTFQNGNGLDLYKAAEELGIEVOREERNNTATQTSIGTQTQATGTERG 180  
 DB 121 Ilyiponydytfeqngnlgldlykaaeelgievoreernnatqtslygtqtaigtterg 180

```

Db 121 llylpqnyqydeggngldlvkaaeajievgreemnatagtslgtlqtaiglteng 180
Oy 181 IYLSAPOIDKLQKACALGSAESIYONANKAKTVLSGIOSIIGSVLAGMDLEALON 240
Db 181 lvsapqldklkqtkagqalgsaesivgnankaktvlsigslgsvlaagmdlealqn 240
Oy 241 NSNGHALAKAGLELTNSLIENIANSVKTLDEFGEIOISQFSKLQNIKGLGTGDKLNKNG 300
Db 241 nsnghalakagletnsleniansvktldfegqisqfsgklgnkgtlgtlqdklnkng 300
Oy 301 GLDKAGLGIDVYISGLSGATLALVADKNASTAKKVGAGFELANOVGNITRAVSSYILA 360
Db 301 gldkaglgldvysglsgataalvldaknastakkvagfelaangvgnitkavssyila 360
Oy 361 QRVAGLSSTGVALIISTVSLAISPLAFAGIADKFNNAKLESEYAEFFKLGIDGDL 420
Db 361 qrvaglsstgvaliistvslaisplafagiadkfnhaksleesyaerfklgydgdl 420
Oy 421 LAEYRGGTIDASYATINTALAAIAGVSAADLTFEKVHNLVITNSKKEKVTIONW 480
Db 421 laeyrggtidasvtaintaalaagvsaadltfekvhnlvitnskkekvtionw 480
Oy 481 FREADFAKEVPYKATKDEKIEEITGONGERTSKOVDDLAKNGKITQDELKRVDMY 540
Db 481 freadfakevpkyatkdekieleitgongertskqvddllakngkitqdelksrvdmny 540
Oy 541 ELKHSKNVTNSLDKLISVSFAFTSSNDSRNVLVAPTSMLOSLSLQFARSOQHSYGL 600
Db 541 ellkhsknvtnsldklisvsfaftsndsrvnlvaptsmldslslqfatsqhsygl 600
Oy 601 RPSGSGQDMSYGLRPGSGSOHWSYGLRPGSGSQDMSYGLRPGSGSQDMSYGLRPGSGSQDW 660
Db 601 rpsgsgqdwsgylrpgsgsqhwsyglrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsqdw 660
Oy 661 SYGLRPGSGSQHWSYGLRPGSGSQDMSYGLRPGGS 695
Db 661 syglrpgsgsqhwsyglrpgsgsqdwsgylrpggs 695

```

RESULT 3  
Y58133  
ID Y58133 standard; Protein: 695 AA.  
XX Y58133;  
AC Y58133;  
XX  
DT 07-MAR-2000 (first entry)  
XX  
De Gonadotropin releasing hormone-leukotoxin fusion protein.  
XX  
KW Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;  
KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;  
KW non-androgenic; steroid; reduction; weight gain; muscle distribution;  
KW fat distribution; male pattern; boar taint; flavour; impalement;  
KW reliable; immunocastration; meat production.  
XX  
OS Chimeric - Mammalia.  
OS Chimeric - Pasteurella haemolytica.  
XX  
FN W09956771-A2.  
XX  
PD 11-NOV-1999.  
XX  
PE 05-MAY-1999; 99WO-CA00360.  
XX  
PR 05-MAY-1998; 98US-0084217.  
XX  
PA (BIOS-) BIOSTAR INC.  
XX  
PI Manns JG, Acres SD, Harland R;  
XX  
DR WPI: 2000-062125/05.  
XX  
DR N-PSDB: 246400.  
XX

```

Pr Production of uncastrated male food animals using vaccines -
XX
XX Claim 22; Fig 3A-3F; 87pp; English.
CC
CC This sequence represents a chimeric gonadotropin releasing
CC hormone (GnRH)-leukotoxin (LKT) fusion protein, which may be
CC used as a vaccine. The LKT portion of the protein acts to enhance
CC the immunogenicity of the multimeric GnRH portion (Y58133). The invention
CC relates to a method of using two GnRH immunogen vaccines to produce
CC uncastrated male animals for meat production, one vaccination prior to
CC or during the fattening period to reduce circulating testosterone levels,
CC and the second vaccination about 2-8 weeks before slaughter to
CC substantially reduce androgenic and/or non-androgenic steroids. The
CC invention is used to produce food animals that exhibit the weight gain
CC and muscle/fat distribution of male animals without the problems
CC associated with male animals. Such problems include "boar taint", a
CC urine-like odour found in cooked meat of uncastrated pigs which is
CC caused by steroids stored in the tissues, and similar flavour
CC impurities in the meat of other intact male animals. The invention is
CC more reliable than prior art immunocastration techniques.
XX
SQ Sequence 695 AA:

```

Query Match 100.0%; Score 3591; DB 21; Length 695;  
Best Local Similarity 100.0%; Pred. No. 2.7e-233;  
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 MATVIDRSQHSYGLRPGSGSQDMSYGLRPGSSOHWSYGLRPGSGSQDMSYGLRPGSGSQ 60
Db 1 matvidrsgqhsyglrpgsgsqdwsgylrpgsgsqhwsyglrpgsgsqdwsgylrpgsgsq 60
Oy 61 HWSYGLRPGSGSQDMSYGLRPGSGSOHWSYGLRPGSGSQDMSYGLRPGSGSPFKTGAKKI 120
Db 61 hwsyglrpgsgsqdwsgylrpgsgsqhwsyglrpgsgsqdwsgylrpgsgsfpkcgakki 120
Oy 121 ILYTPONOYDTEGONGIODLVKAEELGIEVQREERNIAAQNISLGTIONAIGLTENG 180
Db 121 llytpnoydeggngldlvkaaeajievgreemnatagtslgtlqtaiglteng 180
Oy 181 IYLSAPOIDKLQKACALGSAESIYONANKAKTVLSGIOSIIGSVLAGMDLEALON 240
Db 181 lvsapqldklkqtkagqalgsaesivgnankaktvlsigslgsvlaagmdlealqn 240
Oy 241 NSNGHALAKAGLELTNSLIENIANSVKTLDEFGEIOISQFSKLQNIKGLGTGDKLNKNG 300
Db 241 nsnghalakagletnsleniansvktldfegqisqfsgklgnkgtlgtlqdklnkng 300
Oy 301 GLDKAGLGIDVYISGLSGATLALVADKNASTAKKVGAGFELANOVGNITRAVSSYILA 360
Db 301 gldkaglgldvysglsgataalvldaknastakkvagfelaangvgnitkavssyila 360
Oy 361 QRVAGLSSTGVALIISTVSLAISPLAFAGIADKFNNAKLESEYAEFFKLGIDGDL 420
Db 361 qrvaglsstgvaliistvslaisplafagiadkfnhaksleesyaerfklgydgdl 420
Oy 421 LAEYRGGTIDASYATINTALAAIAGVSAADLTFEKVHNLVITNSKKEKVTIONW 480
Db 421 laeyrggtidasvtaintaalaagvsaadltfekvhnlvitnskkekvtionw 480
Oy 481 FREADFAKEVPYKATKDEKIEEITGONGERTSKOVDDLAKNGKITQDELKRVDMY 540
Db 481 freadfakevpkyatkdekieleitgongertskqvddllakngkitqdelksrvdmny 540
Oy 541 ELKHSKNVTNSLDKLISVSFAFTSSNDSRNVLVAPTSMLOSLSLQFARSOQHSYGL 600
Db 541 ellkhsknvtnsldklisvsfaftsndsrvnlvaptsmldslslqfatsqhsygl 600
Oy 601 RPSGSGQDMSYGLRPGSGSOHWSYGLRPGSGSQDMSYGLRPGSGSQDMSYGLRPGSGSQDW 660
Db 601 rpsgsgqdwsgylrpgsgsqhwsyglrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsqdw 660
Oy 661 SYGLRPGSGSQHWSYGLRPGSGSQDMSYGLRPGGS 695

```

Db 661 syglrpgssqhwsgylrpgsqdwsyglrpgss 695

# RESULT 4

ID W79570 standard; Protein: 544 AA.

W79570;

24-DEC-1998 (first entry)

LKT-GnRH chimeric protein.

Chimera: PCB11; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;

cytotoxic activity; antigen presentation; immune response; vaccine; tumour.

Synthetic.

WO9806848-A1.

19-FEB-1998.

08-AUG-1997; 97WO-CA00559.

09-AUG-1996; 96US-0694865.

(UYSA-) UNIV SASKATCHEWAN.

Manns JG, Potter AA.

WPI; 1998-159540/14.

N-PSDB; V61532.

Chimeric protein of leukotoxin and gonadotropin releasing hormone

useful for, e.g. preparation of vaccines for reduction of incidence

of mammary tumours in mammals

Disclosure; Figure 7.1-5; 118pp; English.

The present sequence represents the LKT-GnRH chimeric protein from

PCB11. This plasmid contains the LKT 111 polypeptide fused to

four copies of the GnRH peptide. This chimera lacks cytotoxic activity

which enables there to be an increase in antigen presentation and thus an

optimal immune response. The removal of this region also enables the

truncated LKT to be expressed at much higher levels and allows the amount

of antigen administered to be reduced. This chimeric protein comprises a

leukotoxin polypeptide, several multimers, and a GnRH sequence. The

chimeric protein can be used as a vaccine to help reduce the incidence of

mammary tumours in a mammalian individual.

Sequence 544 AA:

Query Match 73.6%; Score 2642.5; DB 19; Length 544;  
Best Local Similarity 99.4%; Pred. No. 9, 6e-170;  
Matches 534; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 111 SPFKTAKKIIILYIPONTQYDTBOGNGLODLYAAEELGIEVOREERNINATQTSIGTI 170  
DB 8 sfipktakkiillyipnygydteqngldlykaaeelgievreerlnatqtslygtl 67  
QY 171 OFPAIGTEGIVYSAPOIDKLOKTKAGALGSAESIYONANKATVLSGIOSILSYLA 230  
DB 68 gtaigltetgviysapqidljqktkagqalgsaesivqnaakaktvlsyqslisylvla 127  
QY 231 GMDLDEALONNSNQHALAKAGLELTNSLENIANSYKTLDEFEQJTSQFSGSKLQNTKGLG 290  
DB 128 gmdldealgnsnqhalakagletnsleniansvktldefeqsgsklqntkglg 187  
QY 291 TLGDKLKNIGGDLKAGLIGDIVISGLISGATALVLDKKNASTAKKYGAGFEELANQYVGN 350  
TT

Db 188 lfgdklknigldkagligdivisgllsgataalvjadkhnastakkyagfclanqvvgni 247  
QY 351 TKAVSSYIIAQRVAAGLSSGPAVALIASTVSLAISPLAFAGIADKFENAKSLSESYAERF 410  
DB 248 tkavssyiliaqrvaagllssgpaavalaastvslaisplafagladkfennakslesyaeft 307  
QY 411 KILGVDGNLILAEYQRTGTIDASVTAINTAALAIAGVSAANA---DIPEKVKHNLVI 467  
DB 308 kligvgdnlllaeyqrgtgidasvtaintaalaiagvsanaankldlfevkhnlv 367  
QY 468 TNSKKKVTIOMWFERADPFKEVPNKATKDEKIEELIIGQNGERITSKOVDLIAGNK 527  
DB 368 tnskkkvtiomwferadpfkevnpnkatkdekieleligngerltskqvddliakngk 427  
QY 528 ITODELSKYVDNELKHSKNVTNSLDKLISSVSATSSNDSFNVLVAPTMDOSSLSS 587  
DB 428 itodelskvyvdnelkhsknvtnsldkllssvsatssndsfnnvlvaptmldogslss 487  
QY 588 QFARGSQHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPGSS 644  
DB 488 qfargsqhwsyglrpgssqdwsyglrpgssqhwsgylrpgsqdwsyglrpgss 544

## RESULT 5

W03943 ID W03943 standard; Protein: 544 AA.

W03943;

20-NOV-1996 (first entry)

LKT-GnRH protein fusion from PCB11.

Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;

fusion protein; immunogen; vaccine; fertility control;

contraceptive; sterilisation.

Chimeric Pasteurella haemolytica A1 strain B122;

Chimeric synthetic.

Key Location/Qualifiers

Domain 1..493

Domain /label= LKT

Domain /label= 494..544

Domain /label= GnRH\_repeat\_domain

W09624675-A1.

15-AUG-1996.

24-JAN-1996; 96WO-CA00049.

10-FEB-1995; 95US-0387156.

(UYSA-) UNIV SASKATCHEWAN.

Manns JG, Potter AA.

WPI; 1996-384447/38.

N-PSDB; T37177.

Gonadotropin-releasing hormone multimer fusion proteins - with

leukotoxin polypeptide for increased immunogenicity, useful in

antifertility vaccine prodn.

Claim 8: Fig 7A-7E; 87pp; English.

A chimeric protein (W03943) is composed of a fusion between

a truncated leukotoxin (LKT-111) from Pasteurella haemolytica

and a 4-copy gonadotropin-releasing hormone (GnRH) repeat

sequence (see also W03944). It is the product of a chimeric

gene (T37177) produced by deleting an approx. 1300 bp sequence

from PCB113 (see also T37176) coding for amino acids 352-784

CC of LKT-352. Recombinant plasmid pCB111 (LKT 111:4 copy GnrH,  
 CC ATCC 69748) was obt'd. Escherichia coli transformants produced  
 CC the chimeric protein, which is useful as a vaccine for fertility  
 CC control, esp. immunological sterilisation of domestic or farm  
 CC animals.

XX Sequence 544 AA;

Query Match 73.3%; Score 2633.5; DB 17; Length 544;  
 Best Local Similarity 99.3%; Pred. No. 3.9e-169;  
 Matches 533; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

OY 111 SEPKTGAKKIIILYIPQNYDYDEQNGLODLYKAAEELGIEVOREERNNIATAQTSIGTI 170  
 |||||||  
 DB 8 sfprtgakkiilypqnydyteqngldlykkaeeigievqreernniataqtsigti 67  
 |||||||  
 OY 171 QTAIGLTERGIVLSAPQIDKLLQTKAGQALGSAESIYQANAKATVLSGIOSILGSVLA 230  
 |||||||  
 DB 68 qtaiglttergivilsapqidkllqtkagqalgsaesivqnaankaktvlsigilgsvla 127  
 |||||||  
 OY 231 GMDLDEALQNNNOHALKAGLELNSLIENANSVKTLDERGEQISQFSGKLQNIKGLG 290  
 |||||||  
 DB 128 gmdldealqnnsnghalakagletnslienansvktldefgeqisqfsgklqnikglg 187  
 |||||||  
 OY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATAVLADKNASTAKKVGAFELANQVGN 350  
 |||||||  
 DB 188 tlgdklknigldkagldivisgllsgataalvladknastakkvagfelaanqvgni 247  
 |||||||  
 OY 351 TKAVSSYTLAORVAAGLSTGFPVALLASTVSLASPLAFAGIADKFHNAKLSLEYAERF 410  
 |||||||  
 DB 248 tkavssyllaqrvaaglstgfpvaallastvslasplafagiadkfhnaklsesyaerf 307  
 |||||||  
 OY 411 KKLGYDGNLLAEYRGCTIDASTAINTALAIAGVSAAA--DLTEPKVHNHVI 467  
 |||||||  
 DB 308 kklgydgnllaeayrgctidaastaintalaaigvsaanaakldltekvhnhivi 367  
 |||||||  
 OY 468 TNSKKEKXTIONWFEAPFAKPEVPYKATKDEKIEIIGONGERTISROVDLAKKNGK 527  
 |||||||  
 DB 368 tnskkekvtionwfeapfakpevpkyatkdekieleiigongertiskqvdlakngk 427  
 |||||||  
 OY 528 ITODELSKVDNYELLKSKNVTNSLDKLSSVSASTSSNDNRNVLPPTSMLDQSLSL 587  
 |||||||  
 DB 428 itodelskvndnyellkshkvntnsldklssvsastssndnrnlvptsmldqslsl 487  
 |||||||  
 OY 588 OFARSSQHWSTGLRPGSSQDMSTGLRPGSSQDMSTGLRPGSSQDMSTGLRPGSS 644  
 |||||||  
 DB 488 ofarssqhwstglrpgssqdmstglrpgssqdmstglrpgssqdmstglrpgss 544

RESULT 6  
 ID W03942 standard: Protein; 977 AA.

XX W03942;

XX 20-NOV-1996 (first entry)

DE LKT-GnrH protein fusion from pCB113.

KW Leukotoxin; LKT; gonadotropin-releasing hormone; GnrH;  
 fusion protein; immunogen; vaccine; fertility control;

KW contraceptive; sterilisation.

XX Chimeric Pasteurella haemolytica A1 strain B122;

OS Chimeric synthetic.

XX Key location/Qualifiers

FT Domain 1..929  
 FT /label= LKT  
 FT 927..977  
 FT /label= GnrH\_repeat\_domain

PN W09624675-A1.

XX 15-AUG-1996.

XX 24-JAN-1996; 96WO-CA00049.

XX 10-FEB-1995; 95US-0387156.

XX (UySA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1996-384447/38.

XX N-PSDB: T37176.

PT Gonadotropin-releasing hormone multimer fusion proteins - with  
 PT leukotoxin polypeptide for increased immunogenicity, useful in  
 PT antifertility vaccine prodn.

PS Claim 7; Fig 5A-5H; 87pp; English.

CC A chimeric protein (W03942) is composed of a fusion between  
 CC a truncated Leukotoxin (LKT-352) from Pasteurella haemolytica (see  
 CC also W03945) and a 4-copy gonadotrophin-releasing hormone (GnrH)  
 CC repeat sequence (see also W03944). It is the product of a  
 CC chimeric gene (T37176) produced by ligating a synthetic sequence  
 CC for the 4-copy GnrH into vector pAA352 (ATCC 68283), which carries  
 CC the LKT-352 gene. Recombinant plasmid pCB113 (LKT 352:4 copy  
 CC GnrH, ATCC 69749) was obt'd. Escherichia coli transformants  
 CC produced the chimeric protein, which is useful as a vaccine for  
 CC fertility control, esp. immunological sterilisation of  
 CC domestic or farm animals.

XX Sequence 977 AA;

Query Match 67.6%; Score 2426; DB 17; Length 977;  
 Best Local Similarity 55.1%; Pred. No. 7.5e-155;  
 Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps

OY 111 SEPKTGAKKIIILYIPQNYDYDEQNGLODLYKAAEELGIEVOREERNNIATAQTSIGTI 170  
 |||||||  
 DB 8 sfprtgakkiilypqnydyteqngldlykkaeeigievqreernniataqtsigti 67  
 |||||||  
 OY 171 QTAIGLTERGIVLSAPQIDKLLQTKAGQALGSAESIYQANAKATVLSGIOSILGSVLA 230  
 |||||||  
 DB 68 qtaiglttergivilsapqidkllqtkagqalgsaesivqnaankaktvlsigilgsvla 127  
 |||||||  
 OY 231 GMDLDEALQNNNOHALKAGLELNSLIENANSVKTLDERGEQISQFSGKLQNIKGLG 290  
 |||||||  
 DB 128 gmdldealqnnsnghalakagletnslienansvktldefgeqisqfsgklqnikglg 187  
 |||||||  
 OY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATAVLADKNASTAKKVGAFELANQVGN 350  
 |||||||  
 DB 188 tlgdklknigldkagldivisgllsgataalvladknastakkvagfelaanqvgni 247  
 |||||||  
 OY 351 TKAVSSYTLAORVAAGLSTGFPVALLASTVSLASPLAFAGIADKFHNAKLSLEYAERF 410  
 |||||||  
 DB 248 tkavssyllaqrvaaglstgfpvaallastvslasplafagiadkfhnaklsesyaerf 307  
 |||||||  
 OY 411 KKLGYDGNLLAEYRGCTIDASTAINTALAIAGVSAAA----- 454  
 |||||||  
 DB 368 gvlstlqysqamfehanklnkiveweknbnhkyfengydarylanqdmkfln 427  
 |||||||  
 OY 455 ----- 454  
 |||||||  
 DB 428 lnkelaervalitqgwdnlnlgdlaglsrlgekvlsgkayvdafegkhkacklvgld 487  
 |||||||  
 OY 455 ----- 454

Db 488 sanglidsnsgkakqhllfrp1ltpgrehervgtkyeyltnlrsvdskltdg 547  
 QY 455 ----- 454  
 Db 548 aasftldlnvgrlgeldnagnvltkktliaklgdgnvfvgsgttdidggeyd 607  
 QY 455 ----- 454  
 Db 608 rvhysrgnygaltdatketegsgtyvnrfvetgkalhevtsthtalvgnreekieyrhs 667  
 QY 455 ----- 454  
 Db 668 nqghagyytkdtkaveeiigtshndifkyskfndafnggdvdtidngdnrlfgyk 727  
 QY 455 ----- 454  
 Db 728 gddldgngddfidggkgnldlhggkgddlfvhrkygdndltdsgndklfsdsnlk 787  
 QY 455 DLFEKVKHNLVTNKKKEKVTIQQNMFREADFAKEVNPYKATDEKIEITIGONGERITS 514  
 Db 788 dlfevkhnvltnskkevtlgnwtfreadfakvnpnykatdekleeiigngerits 847  
 QY 515 KQYDDLIAKNGKITQDELKRVNDNTLLKHSKNVTNSLDKLISVSAPFTSSNDNRNVLV 574  
 Db 848 kyvddliakngkltqdeliskvndnyellkhsknvtnslldklisvsafesndsrnvlv 907  
 QY 575 APTSMIDQSSIOFARGSOHWSTGLRPGSGSDMWGLRPGSSOHWSTGLRPGSGSD 634  
 Db 908 apsmidqssiqfargsgqhwstglrpgsgsdwsgylrpgsgsqhwstglrpgsgsd 967  
 QY 635 WSYGLRPGGS 644  
 Db 968 wsyglrpggs 977  
 RESULT 7  
 ID W79569 standard; Protein; 977 AA.  
 AC W79569;  
 DT 24-DEC-1998 (first entry)  
 DE LKT-GnRH chimeric protein.  
 KW Chimera; pcB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;  
 KW cytotoxic activity; antigen presentation; immune response; vaccine;  
 tumor.  
 OS Synthetic.  
 PN WO9806848-A1.  
 PD 19-FEB-1998.  
 PF 08-AUG-1997; 97WO-CA00559.  
 PR 09-AUG-1996; 96US-0694865.  
 PA (UYSA-) UNIV SASKATCHEWAN.  
 PI Manns JG, Potter AA;  
 DR WPI; 1998-159540/14.  
 DR N-PSDB; V61531.  
 PT Chimeric protein of leukotoxin and gonadotropin releasing hormone  
 PT useful for, e.g. preparation of vaccines for reduction of incidence  
 PT of mammary tumours in mammals  
 PS Disclosure; Figure 5.1-8; 118pp; English.

CC The present sequence represents the LKT-GnRH chimeric protein from  
 CC pcB113. This plasmid contains the LKT 352 polypeptide (W79568) fused to  
 CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity  
 CC which enables there to be an increase in antigen presentation and thus an  
 CC optimal immune response. The removal of this region also enables the  
 CC truncated LKT to be expressed at much higher levels and allows the amount  
 CC of antigen administered to be reduced. This chimeric protein comprises a  
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
 CC chimeric protein can be used as a vaccine to help reduce the incidence of  
 CC mammary tumours in a mammalian individual.  
 SQ Sequence 977 AA;  
 Query Match 67.6%; Score 2426; DB 19; Length 977;  
 Best Local Similarity 55.1%; Pred. No. 7.5e-155;  
 Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;  
 QY 111 SPFKTGAKKILLYIPONYOYDFEONGLODLVKAABELGIEVOREBRNNIATQTSIGTI 170  
 Db 8 sfktgkakkillyipnydydegngldlvkaaeelgievgreennlatatqstlgtl 67  
 QY 171 QTAIGLTERGIVSAPQIDKLOKTKAGALGSABESIVONANKAKTVLSGISTISVLA 230  
 Db 68 qlaiglttergvisapqidlklqtkagaglsaeslvgnankakvlsigilgsvla 127  
 QY 231 GMDLDEALONNSNOHALAKAGLELTNSLIENANSVKTIDFEFEOISOPCSKLONTKGLC 290  
 Db 128 gmdldealnnsnqhalaakagletnslienansvktidefegqisqfysknglkylg 187  
 QY 291 TLGDKLKNIGGLDKAGIGLDIVISGLSGATALVLDKRNASTKAKVAGFELANOVVNT 350  
 Db 188 tlgdklknigldkagigldivisglsgataalvldknaastakvgaefelanovvnt 247  
 QY 351 TKAVSSYTLAQRVAAGISSGPVAALIASVSLAIPFAAGIADKFNHAKKSLESVAER 410  
 Db 248 tkavssytlagrvaagissgpvaaliasvslaisplataidkfnhakslesyaer 307  
 QY 411 KRLGDDNLLAEYORGTGTIDASVTAINPALAIGVSAANA----- 454  
 Db 308 krlgddnllaeqrgtgidasvtaintalaaigvsaanaagvsiallvsqtl 367  
 QY 455 ----- 454  
 Db 368 gvlstllqyskgamfehvanklnhkiwelnhngknyfengydarylanlgdmkflln 427  
 QY 455 ----- 454  
 Db 428 lnkelaervialtqgqwdnnigdlaglsrlgkvlsgkayvdafeegkhikadklvqld 487  
 QY 455 ----- 454  
 Db 488 sanglidsnsgkakqhllfrp1ltpgrehervgtkyeyltnlrsvdskltdg 547  
 QY 455 ----- 454  
 Db 548 aasftldlnvgrlgeldnagnvltkktliaklgdgnvfvgsgttdidggeyd 607  
 QY 455 ----- 454  
 Db 608 rvhysrgnygaltdatketegsgtyvnrfvetgkalhevtsthtalvgnreekieyrhs 667  
 QY 455 ----- 454  
 Db 668 nqghagyytkdtkaveeiigtshndifkyskfndafnggdvdtidngdnrlfgyk 727  
 QY 455 ----- 454  
 Db 728 gddldgngddfidggkgnldlhggkgddlfvhrkygdndltdsgndklfsdsnlk 787  
 QY 455 DLFEKVKHNLVTNKKKEKVTIQQNMFREADFAKEVNPYKATDEKIEITIGONGERITS 514  
 Db 788 dlfevkhnvltnskkevtlgnwtfreadfakvnpnykatdekleeiigngerits 847



```

QY 515 KOVDLLAKNGKXITODELSKVNDVELLKHKNVTNSLDKLISVSAPFTSSNDSRNVLV 574
DB 848 KQVDDLLAKNGKXITODELSKVNDVELLKHKNVTNSLDKLISVSAPFTSSNDSRNVLV 907
QY 575 APTSMLODLSLLOFANGSQHMSYGLRPGSGSODMSYGLRPGSSQHMYSYGLRPGSGSOD 634
DB 908 APTSMLODLSLLOFANGSQHMSYGLRPGSGSODMSYGLRPGSSQHMYSYGLRPGSGSOD 967
QY 635 WSYGLRPGGS 644
DB 968 WSYGLRPGGS 977

RESULT 8
Y33929
ID Y33929 standard; Protein; 490 AA.
XX
AC Y33929;
XX
DT 09-NOV-1999 (first entry)
XX
DE Leukotoxin carrier protein (LKT 114).
XX
KM Myostatin: mouse; rabbit; human; baboon; bovine; porcine; ovine; chick;
KM turkey; zebrafish; immune response; vaccine; body weight; muscle mass;
KM mammary gland tissue; lactation; feed uptake; muscle degeneration;
KM GDF11 activity; LKT 114; leukotoxin carrier protein.
XX
OS Pasteurella haemolytica.
XX
PN WO9942573-A1.
XX
PD 26-AUG-1999.
XX
PF 19-FEB-1999; 99WO-CA00128.
XX
PR 19-FEB-1998; 98US-0075213.
XX
PA (BIOS-) BIOSTAR INC.
XX
PI Barker CA, Morsey M;
XX
DR WPI: 1999-527471/44.
XX
N-PSDB: X99361.
XX
PT New myostatin peptide, multimers and immunocjugates for eliciting
XX an immune response in a vertebrate against a myostatin immunogen
XX
PS Example 2; Fig 15; 109pp; English.
XX
CC The invention provides myostatin peptides consisting of 3-100 amino
CC acids, derived from a region of mouse, rabbit, human, baboon, bovine,
CC porcine, ovine, chick, turkey or zebrafish myostatin (see sequences
CC Y33930-939). The myostatin peptides are derived preferably from a region
CC of amino acid residues 1-275, 25-300, 50-325 or 75-350 of the above
CC sequences. The peptides and the nucleic acids encoding the peptides are
CC useful as vaccines for eliciting an immune response in a vertebrate
CC against a myostatin immunogen. They result in increasing body weight,
CC muscle mass, number and size of muscle cells, muscle strength, mammary
CC gland tissue, lactation, appetite or feed uptake, life span of the
CC vertebrate, and cause a reduction in body fat content, useful for muscle
CC wasting conditions. The vaccines are also useful for treating a disorder
CC which comprises degeneration or wasting of muscle in a vertebrate, and
CC useful for modulating GDF11 activity. The present sequence represents a
CC leukotoxin carrier protein (LKT 114) which is used as an immunological
CC carrier protein to form a myostatin immunocjugate.
XX
SQ Sequence 490 AA:

```

Query Match 65.6%; Score 2354; DB 20; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 2e-150;

```

Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 111 SFPRTGAKKILITIPQNTQYDTQNGNGLODLYKAAEELGIEVQREERNNTTAQTSIGTI 170
DB 8 SFPKCGAKKILITIPQNYQYDTQNGNGLODLYKAAEELGIEVQREERNNTTAQTSIGTI 67
QY 171 QTAIGLTFRGIVLSAPQIDKLOKTKAQALGSAESIYQNNKAKTVLSGIQSIIGSVLA 230
DB 68 QTAIGLTFRGIVLSAPQIDKLOKTKAQALGSAESIYQNNKAKTVLSGIQSIIGSVLA 127
QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDERFGEQISQSGKLQNIKGLG 290
DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDERFGEQISQSGKLQNIKGLG 187
QY 291 TLGDKLKNIIGLDKAGLDVYISGLSGATPAALVADKNASTAKKVGAGFELANQVGN 350
DB 188 TLGDKLKNIIGLDKAGLDVYISGLSGATPAALVADKNASTAKKVGAGFELANQVGN 247
QY 351 TKAVSSYTLAQRVAAAGLSSTGPVAAALIASTVSLAISPAPAGIADKFNHAKSLESYAERF 410
DB 248 TKAVSSYTLAQRVAAAGLSSTGPVAAALIASTVSLAISPAPAGIADKFNHAKSLESYAERF 307
QY 411 KKLGYDGNLAEYQRTGTIDASVTATNTALAAIAGCVSAAADLPFEKYKHNLYITNS 470
DB 308 KKLGYDGNLAEYQRTGTIDASVTATNTALAAIAGCVSAAADLPFEKYKHNLYITNS 367
QY 471 KKEKVTIONMFREADFAKEVFNKATKDEKTEIEITIGNGERTSKQVDLLAKNGKXITQ 530
DB 368 KKEKVTIONMFREADFAKEVFNKATKDEKTEIEITIGNGERTSKQVDLLAKNGKXITQ 427
QY 531 DELSKVVDNVELLKHKNVTNSLDKLISVSAPFTSSNDSRNVLVAPTSMLODLSLLOFA 590
DB 428 DELSKVVDNVELLKHKNVTNSLDKLISVSAPFTSSNDSRNVLVAPTSMLODLSLLOFA 487
QY 591 RGS 593
DB 488 RGS 490

RESULT 9
R34547
ID R34547 standard; Protein; 936 AA.
XX
AC R34547;
XX
DT 23-AUG-1993 (first entry)
XX
DE GNRH-leukotoxin gene fusion prod.
XX
KM Vector: LKT 352; flanking; recombinant; antigen; somatostatin;
KM gonadotropin releasing hormone; rotavirus viral protein 4;
KM carrier protein; lactation; reproduction; SRIF.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Protein 1..926
FH Peptide /note="recombinant leukotoxin protein"
FH Peptide 927..936
FH /note="GNRH"
XX
PN WO9308290-A.
XX
PD 29-APR-1993.
XX
PF 15-OCT-1992; 92WO-CA00449.
XX
PR 16-OCT-1991; 91US-0779171.
XX
PR 14-OCT-1992; 92US-0960932.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Hughes HPA, Potter AA, Redmond MJ;

```

XX WPI: 1993-152482/18.  
 DR N-PSDB: Q41322.  
 XX  
 PT Immunological carrier system with enhanced immunogenicity -  
 PT comprises chimeric protein comprising leuco:toxin peptide or  
 PT homologous protein fused to antigen esp. somatostatin or  
 PT gonadotropin releasing hormone  
 XX  
 PS Example 2; Fig 8; 95pp; English.  
 XX  
 CC Oligonucleotides contg. sequences from bovine gonadotrophin  
 CC releasing hormone (GnRH) gene were constructed on a Pharmacia Gene  
 CC Assembler using standard phosphoramidite chemistry. The oligo-  
 CC nucleotides were annealed and ligated into vector pAA352 (contg.  
 CC the Pateurella haemolytica leuco-toxin gene) which had been digested  
 CC with BamHI. The ligated DNA was used to transform E. coli strain  
 CC MH3000. Transformants contg. the oligonucleotide inserts were  
 CC identified by restriction endonuclease mapping and the recombinant  
 CC plasmid designated pAA502. The chimeric protein produced from the  
 CC plasmid works to bring about a larger immune response than the antigen  
 CC alone, i.e. the leukotoxin works as a carrier protein.  
 CC See also R34545-8.  
 XX  
 XX Sequence 936 AA:

Query Match 60.5%; Score 2173; DB 14; Length 936;  
 Best Local Similarity 52.9%; Pred. No. 6,9e-138;  
 Matches 491; Conservative 0; Mismatches 2; Indels 436; Gaps 1;

QY 111 SPFTGAKIILYIPONYQYDEQNGLODLVKAELGIEVORENNNTATQTSIGTI 170  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 8 sfptgakkililypnyqydeqngldvlkaaeelgievorenmatatqtsigtl 67  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 171 QTAIGTFERGIVASAPQIDKLOKTRAGALGSAESTIVONANKAKTYLGSITLSVLA 230  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 68 qtaigttergivilsapqidlilqktkagalgisaesivqnaankaktylsgisllsvla 127  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 231 GMDLDAFALNNSQNHAKLAKAGELTNSLTENIANSVKTLDPEFEOISQFESKTONKIG 290  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 128 gmdldealqnsnqhalakageltnslteniansvktldexceqtsqfsgklqnkylg 187  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 291 TLGDCKRNITGDLKAGLGLDVISGLSGATPAALVLDKNASTAKKYGAGFELANQVGN 350  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 188 tlgdcklknigylkdagldvlsqllsgataalvldknaastakkygagfelanqvgnl 247  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 351 TKAVSSYIIAQRVAAGLSTGTPVAALIASTVSLAISPLAFAGIADKFNNAKSLSEIAERF 410  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 248 tkavssyilaqrvaaglststgvaalaaastvslaisplafagidkfnhakslesyerf 307  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 411 KKLGYGDNLLAEYRGSTGTIDASVTAINTALAIAGVSAANA----- 454  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 308 kklgygdnlleayrgrtgidasvtalntalaatagvsaanaaasviallvsqilt 367  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 455 ----- 454  
 DB 368 gvastllqyskqamfehvancklnkiveweknbnknyfengydarylanlqdmnkfiln 427  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 455 ----- 454  
 DB 428 lnkelaervialtggqwdnlnlgdlaglsrlgskvlskayvdafeegkhikadklvql 487  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 455 ----- 454  
 DB 488 sanglidsvnsqkacqhlfrtptlltpgtehrerwtgkyeyitklnlnrvdswkilt 547  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 455 ----- 454  
 DB 548 aasstfdlcnvgrjgieldnagnvktketkiaklgegdanrvfsgteldggedgyd 607  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 455 ----- 454

DB 608 rvhysrgnygallidatkebegsyfwnrfvelgkahlvetslhalvgnreeklgyrns 667  
 QY 455 ----- 454  
 DB 668 nngnhagyytkdtkaveeilgtshndifkyskfndafnggdvdtldgndgnrlfsgk 727  
 QY 455 ----- 454  
 DB 728 gddldgngddfidgkgndllhggkgddlfvhrkgdgnlldtsdgnkdksfsdsnlk 787  
 QY 455 DLTFEKVKNHVLVTNKSKEKVTIQNMFREADPAKEVPNPKATYDEKIEELIGNGERITS 514  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 788 dltfekvknhlvtlnskkdvltlgnwfreaddakevnykatkdekleeilgnggerits 847  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 515 KOYDDLIAGNGKITODELSKVVDNELKHSKNVTNSLDKLISSVSAFSSNDSSRNVLV 574  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 848 kyvddliakngkitqdeliskvvdnyellkhsnvnnsldkllssvsafssndssrnvlv 907  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 575 APTSMLOSLSSLOFARSGQHSYGLRPG 603  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 908 aptsmldgslsslqfarsgqhsyglrpg 936  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 10  
 R52748  
 ID R52748 standard; Protein; 1069 AA.  
 XX  
 AC R52748;  
 XX  
 DT 01-JUL-1994 (first entry)  
 DE  
 XX Bovine IFNgamma/LTK chimeric protein encoded by plasmid pAA497.  
 DE  
 XX Bovine: Interleukin-2; IL2; P. haemolytica; leukotoxin; LTK; IFN.  
 KW ltkA; chromosome walking; fusion protein; vaccine; interferon; gamma;  
 KW monoclonal; polyclonal; antibody.  
 XX  
 OS Pasteurella haemolytica - chimera.  
 OS Bos taurus - chimera.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..926  
 FT /note= "Recombinant leukotoxin peptide [splic]"  
 FT Peptide 927..1069  
 FT /note= "Bovine IFNgamma"  
 XX  
 PN US5273889-A.  
 XX  
 PD 28-DEC-1993.  
 XX  
 PF 22-AUG-1990; 90US-0571301.  
 XX  
 PR 22-AUG-1990; 90US-0571301.  
 PR 16-OCT-1991; 91US-0777715.  
 XX  
 PA (CIBA ) CIBA GEIGY CANADA LTD.  
 PA (UYSA-) UNIV SASKATCHEWAN.  
 XX  
 PI Campos M, Hughes HPA, Potter A;  
 XX  
 DR WPI: 1994-006687/01.  
 DR N-PSDB: Q54213.  
 XX  
 PT Immunogenic fusion proteins of gamma-interferon and immunogenic  
 PT leukotoxin - used in vaccines and to raise monoclonal and polyclonal  
 PT antibodies  
 XX  
 PS Disclosure; Fig 7; 56pp; English.  
 XX  
 CC This sequence represents a fusion between bovine gamma interferon  
 CC (IFNgamma) and Pasteurella haemolytica leukotoxin (LTK). The  
 CC leukotoxin gene, ltkA, was isolated from a gene library of P.  
 CC haemolytica by chromosome walking. Immunogenic fusion proteins

CC such as this can be used in vaccine compositions. It can also be  
 CC used to raise mono- and polyclonal antibodies.

SO Sequence 1069 AA;

Query Match 59.3%; Score 2131; DB 15; Length 1069;  
 Best Local Similarity 52.6%; Pred. No. 5.5e-135;  
 Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

OY 111 SFPKGAKKIILYIPONTQYDTEGNGLODLVKAABELGIEVQREERNNIATAQTSIGTI 170
DB 8 sfptgakkililyipqnygdtdegngldlvkaaeelgievqreernniataqtslgti 67
OY 171 QTAIGLTERGIVLSAPODKLLQKTKAGALGSAESTYQONANKATVLSIGISIGSVLA 230
DB 68 qtaiglttergivilsapqldkllqtkkagalgasaesivqnaankaktvlsigisilgsvla 127
OY 231 GMDLDEALQNNNSOHALKAGLELTNSLENIANSVKTTDEFGEOISQFSKLNKIGLG 290
DB 128 gmdldealqnnnsqhalakagletnsleniensvktldefgesqfsklnqtkgig 187
OY 291 TLGDKLKNIGSIDKAGLGIDVISGLSGATAALVLADKNASTAKRVGAGFELANOVGN 350
DB 188 tlgdklknigsidkaglgldvisgllsgataalvladknastakrvagagfelanqvgn 247
OY 351 TRKAVSYLLAQRVAGLSSTGPAALLASTVSLAISPLAFACIADKFNHAKLESTAEF 410
DB 248 tkavssyllaqrvaaaglsstgpvaallastvslaisplafaciadkfnhakslesyaerf 307
OY 411 KKLGYDGNLLAEYORGTIDASVTATNTALATAGVSAANA----- 454
DB 308 kklgydgnllaeaygrgtdltaavtaintalaalagvsaanaasvlaspiallvsgit 367
OY 455 ----- 454
DB 368 gvlstllgysqamfehvanklhnkiveweknbnghknyfengydarylanlqdmkflin 427
OY 455 ----- 454
DB 428 lnkelaervlaiaiqgwdnlnlagnlsrlgekvlsqaydaleegkhlkadhvlqld 487
OY 455 ----- 454
DB 488 sanglidsvnsngaktqhlflrtpllpptehrvvgkyeyiklnlnvdsukltdg 547
OY 455 ----- 454
DB 548 aasstfdltnvvrfiglndnagntckekkliaqlgeddnvfvgstteidsggyd 607
OY 455 ----- 454
DB 608 rvhsrnygaltdacketeqgsytlvnrfvetgkahlveisthalygnreekleyrhs 667
OY 455 ----- 454
DB 668 nqnhaagrytkdtklaveeilgtsmndifkgsfndafrngddgvdtlqndndrlfpgk 727
OY 455 ----- 454
DB 728 gddlldgngddflidgqgndllbhgkgddflvhrkgdgnndltdsdgndklisfedsnlk 787
OY 455 ----- 454
DB 788 dltefkvhnvltmskkekvtlqmwfpreadpafakvnpnykatkdkiieitigoneerits 514
DB 847 dltefkvhnvltmskkekvtlqmwfpreadpafakvnpnykatkdkiieitigoneerits 847
OY 515 KOVDLIAKGNKTIQDELSEKVVNDVEYELLSKNTNSLKLISVSAFTSSNDSDRNVLV 574
DB 848 kvvdellakngkltqdelsekvvdnyellhsknvtnsldklisvsaftssndsdrrnvlv 907
OY 575 APTSMLDOSLSLOPARCSQ 594
DB 908 aptsmldqslslsfarqsg 927

```

```

RESULT 11
ID W13867
XX W13867 standard; Protein; 1069 AA.
AC W13867;
XX 12-MAR-1997 (first entry)
DT 12-MAR-1997 (first entry)
XX Chimeric protein #2.
DE RTX cytotoxin; cytokine; immunogen; chimeric protein; cytokine; vaccine;
XX Interleukin-2, IL-2; gamma interferon; gamma IFN; leukotoxin; pneumonia;
KW Pasteurella haemolytica; LKT352; respiratory disease; shipping fever;
KM fibrinous pneumonia; cattle; therapy.
XX Synthetic.
OS US5594107-A.
XX 14-JAN-1997.
XX 22-AUG-1990; 90US-0571301.
XX 20-DEC-1993; 93US-0170126.
XX 22-AUG-1990; 90US-0571301.
XX 16-OCT-1991; 91US-0777715.
PA (CIBA ) CIBA GEIGY CANADA LTD.
PA (UUSA-) UNITV SASKATCHEWAN.
XX Campos M, Hughes HPA, Potter A;
XX WPI; 1997-099529/09.
XX DR N-PSDB; T60033.
XX PT Immunogenic chimeric proteins comprising cytokine linked to RTX
XX toxin - useful in vaccines, esp. against shipping fever in cattle
XX Claim 13; Column 37-46; 56pp; English.
XX W13866 and W13867 represent immunogenic chimeric proteins of the
CC invention. This sequence represents a chimeric protein containing the
CC bovine gamma interferon (gamma IFN) sequence and a leukotoxin sequence.
CC The chimeric proteins of the invention comprise a cytokine, selected from
CC Interleukin-2 (IL-2) and gamma IFN, linked to at least one RTX toxin
CC epitope (preferably the sequence shown in W13865). The RTX toxin used to
CC provide the epitope sequence is preferably a leukotoxin, especially the
CC full-length Pasteurella haemolytica leukotoxin. Alternatively, the
CC leukotoxin is a truncated leukotoxin lacking leukotoxic activity,
CC especially LKT352. The chimeric proteins can be used for the production
CC of vaccines against respiratory diseases such as pneumonia, particularly
CC fibrinous pneumonia caused by P. haemolytica, including shipping fever in
CC cattle.
XX SO
XX Sequence 1069 AA;

```

Query Match 59.3%; Score 2131; DB 18; Length 1069;  
 Best Local Similarity 52.6%; Pred. No. 5.5e-135;  
 Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

OY 111 SFPKGAKKIILYIPONTQYDTEGNGLODLVKAABELGIEVQREERNNIATAQTSIGTI 170
DB 8 sfptgakkililyipqnygdtdegngldlvkaaeelgievqreernniataqtslgti 67
OY 171 QTAIGLTERGIVLSAPODKLLQKTKAGALGSAESTYQONANKATVLSIGISIGSVLA 230
DB 68 qtaiglttergivilsapqldkllqtkkagalgasaesivqnaankaktvlsigisilgsvla 127
OY 231 GMDLDEALQNNNSOHALKAGLELTNSLENIANSVKTTDEFGEOISQFSKLNKIGLG 290

```

```

Db 128 gmdldeaiqnnsnqhalakagletnsljenlansvktldefeqisqfsgsklgnlkylg 187
OY 291 TLGDKLKNIGLDKAGLDVTSGLSGATPAALVLADKNASTKRVGAGFELANOVVGN 350
Db 188 tlqgdklknlgldkagldvtsglsgataalvladknastrakvvgagfelanqvvgnl 247
OY 351 TKAVSSYIIAQRVAAGLSSTGPVAALIASTVSLAISPFAIADKFNNAKSLESAERF 410
Db 248 tkavssyIIagrvaaglsstgpvaalIastvslaisplafagldkfnhakslesyaerf 307
OY 411 KRLGYGDNILAEYRGCTGTIDASVTAINATALAINGVSAAA-----454
Db 308 krlgygdnllaeYgrgtldasvtaIntalaaIagvsaagaavslasplallvsglt 367
OY 455 -----454
Db 368 gvtstllgyskgamfehvanklnkIveweknmhknfyfengydarylanldnmkflIn 427
OY 455 -----454
Db 428 lnkqlgaerviaItqgwdnlgldaglsrlgkvisgkayvdafeegkhikadklvql 487
OY 455 -----454
Db 488 saangIIdvsnsgaktqhlftfpIltptgthervygtkyeylklInrvdswkItlg 547
OY 455 -----454
Db 548 aaastldlcnvgrIsIeldnagnvktketklIaklgagdanvfvgstleldgegyd 607
OY 455 -----454
Db 608 rvhysrgnygaltIdatketegsylvnrfvetgkalhevsthtalvgnreekIeyrha 667
OY 455 -----454
Db 668 nqghagyyktclIkaveelIgtshndIfkyskfndaIngdgvdIdgndgnrlfgyk 727
OY 455 -----454
Db 728 gddIldgngddIdfgqgkgnldlhgqkgddIfvhrkygdndIIdtdgndklsfsdnk 787
OY 455 DLTFEKVKNLVTNKKKEKVTIONMFREADFAKEVPNKATFDEKIEELIGONGERITS 514
Db 788 dltfekvkhnlvItnskkekvtIqmwtfreadfakvpenykatdkeIeelIqngertIs 847
OY 515 KQYVDLIANGNGKITODELSKVVDNVELLKHKNVTNSLDKLISVSATFSSWDSRWLV 574
Db 848 kqvddIlaKngkItqdeIstkvvdnveIlkhskvntnsIdklIssvsatfssndsInvlv 907
OY 575 APTSMIDQSLSLQFARGSQ 594
Db 908 aptsmIdqslslsqfargsq 927

```

RESULT 12  
R14482  
R14482 standard; Protein; 926 AA.

XX AC R14482;  
XX DT 15-JAN-1992 (first entry)  
XX DE LKT352.  
XX KW Antigen; leukotoxin; vaccine; Ikta.  
XX OS Pasteurella haemolytica.  
XX PN WO9115237-A.  
XX PD 17-OCT-1991.

```

PF 17-OCT-1991; 91WO-CA00170.
XX
PR 05-APR-1990; 90US-0504850.
XX
PA (UTSA-) UNIV SASKATCHEWAN.
XX
PI Acres SD, Barluk LA, Potter AA, Lawman MDP;
DR WPI; 1991-324967/44.
XX
PT Vaccines for Pasteurella haemolytica infection in cattle -
PT comprise sub-unit antigens from P haemolytica fimbrial protein,
PT plasmid receptor, 50 K outer membrane protein and leukotoxin.
XX
PS Disclosure; Fig 5; 92pp; English.
XX
CC LKT352 is 98% homologous with authentic leukotoxin and migrates
CC to the same position on gels.
CC The LKT352 gene was prep'd. as follows: Ikta, an MaeI fragment
CC contg. the gene was ligated into the SmaI site of pUC13 to form
CC pAA179. From this, two constructs were made in the pUC-based
CC vector, pGH432:laci digested with SmaI. One, pAA342, consisted of
CC the entire MaeI fragment from Ikta while the other, pAA345, contained
CC leukotoxin peptide at high levels while pAA345 expressed full
CC length leukotoxin at very low levels. The 3' end of the Ikta gene
CC of pAA345 was therefore ligated to StyI/BamHI digested pAA342 to
CC yield pAA352 contg. the LKT352 sequence. The protein expressed
CC from the vector can be used to prepare a subunit vaccine with
CC other P. haemolytica antigens, e.g. fimbrial protein, plasmid
CC receptor or 50K outer membrane protein. The vaccines can be used
CC to protect cattle from respiratory diseases such as pneumonia, esp.
CC shipping fever pneumonia.
XX See also R14481, 83,84 and 85.
XX
SQ Sequence 926 AA:

```

Query Match 59.2%; Score 2126; DB 12; Length 926;  
Best Local Similarity 52.6%; Pred. No. 9.7e-135;  
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

OY 111 SFPKTAGKILIIYPONYOYDEQNGLODLVKAABELIEVOREBRNNIATAQTSIGTI 170
Db 8 sfpktagkIilYipnyqydtdegngldlvkaaelgievgreennlatatqtsIgtI 67
OY 171 QTAIGLTGRTIVSAPOIDKLQKTRAGALGSABEIVQNAKAKTVLSIGISLVIA 230
Db 68 qtaIgtltergIvIsapqIdklIqktkagalaGsaeIvqnaakaklvlsIqIsIlgsvIa 127
OY 231 GMDLDEALONNSQOHALAKAGLELTSLJENINSVKTLDEFEQOISOPGSKLQNTKGLG 290
Db 128 gmdldeaiqnnsnqhalakagletnsljenlansvktldefeqisqfsgsklgnlkylg 187
OY 291 TLGDKLKNIGLDKAGLDVTSGLSGATPAALVLADKNASTKRVGAGFELANOVVGN 350
Db 188 tlqgdklknlgldkagldvtsglsgataalvladknastrakvvgagfelanqvvgnl 247
OY 351 TKAVSSYIIAQRVAAGLSSTGPVAALIASTVSLAISPFAIADKFNNAKSLESAERF 410
Db 248 tkavssyIIagrvaaglsstgpvaalIastvslaisplafagldkfnhakslesyaerf 307
OY 411 KRLGYGDNILAEYRGCTGTIDASVTAINATALAINGVSAAA-----454
Db 308 krlgygdnllaeYgrgtldasvtaIntalaaIagvsaagaavslasplallvsglt 367
OY 455 -----454
Db 368 gvtstllgyskgamfehvanklnkIveweknmhknfyfengydarylanldnmkflIn 427
OY 455 -----454
Db 428 lnkqlgaerviaItqgwdnlgldaglsrlgkvisgkayvdafeegkhikadklvql 487

```

```

OY 455 ----- 454
DB 488 sangliidvnsqkaktqhl1ftprlltpgtehrervgtqkyeyicklnnrdswwkltgd 547
OY 455 ----- 454
DB 548 aasstfdlnvvgrigieldnagnvktckekliaklggeddnvfvsgsteldggedyd 607
OY 455 ----- 454
DB 608 rvhysrgnygallidatketeggyfvrnfveclgkallhevtsthalvgnreekieyrs 667
OY 455 ----- 454
DB 668 mngnhagyytktlkaveeligtshndifkyskfndafnggdvdlidngdnrlfgyk 727
OY 455 ----- 454
DB 728 gddldgngddfidgqkgndllhggkgddlfvhrkgdgnlldsdgndklsfsdnlk 787
OY 455 ----- 454
DB 788 dlfevkhmlvltnskkekvtlqmwreadfakvpykatkdekleeilgqngerits 847
OY 515 KOVDLIAKNGKITODELSKVVDNVELLHSHKNVTNSLDKLISSVSASFSSNDSENVLV 574
DB 848 kvyddliakngkltgdelskvvdnyellhshknvnsldkllssvatsfssndsrnvl 907
OY 575 APTSMLOSLSLQFARGS 593
DB 908 aptsmldgslslqfargs 926

```

## RESULT 13

```

R34545 ID R34545 standard; Protein: 926 AA.
AC R34545:
DT 23-AUG-1993 (first entry)
DE Leukotoxin 352 produced from PAA352.
KW Vector: LKT 352; flanking; recombinant; antigen; somatostatin;
KW gonadotropin releasing hormone; rotavirus viral protein 4;
KW carrier protein; lactation; reproduction.
OS Pasteurella haemolytica.
PN WO9308290-A.
PD 29-APR-1993.
PE 15-OCT-1992; 92MO-CA00449.
PR 16-OCT-1991; 91US-0779171.
PR 14-OCT-1992; 92US-0960932.
XX (UYSA-) UNIV SASKATCHEWAN.
PI Hughes HPA, Potter AA, Redmond MJ;
DR N-PSDB: Q41317.
XX WPI: 1993-152482/18.
XX Immunological carrier system with enhanced immunogenicity -
XX comprises chimeric protein comprising leucotoxin peptide or
XX homologous protein fused to antigen esp. somatostatin or
XX gonadotropin releasing hormone
XX Disclosure; Fig 3; 95pp; English.
XX

```

CC Gene libraries of P. haemolytica A1 (strain B122) were constructed  
 CC in lambda gtl1 and pUC13. Resulting clones were used to transform E.  
 CC coli and individual colonies were pooled and screened for reaction  
 CC with serum from a calf which had survived a P. haemolytica infection  
 CC and that had been boosted with a conc. culture supernatant of P.  
 CC haemolytica to increase anti-leukotoxin antibody levels. Positive  
 CC colonies were screened for their ability to produce leukotoxin by  
 CC incubating cell lysates with bovine neutrophils and measuring the  
 CC release of lactate dehydrogenase from the neutrophils. A 4kb  
 CC fragment was obtd. Progressively larger clones were isolated by  
 CC chromosome walking to isolate full length recombinants of ca. 8kb,  
 CC in pAA14. The clone was subjected to restriction enzyme digestion  
 CC to yield two clones, one expressing truncated leukotoxin peptide at  
 CC high levels and the other expressing the full length leukotoxin at  
 CC low levels. The 3' end of the lktA gene from the full length clone  
 CC was ligated to the truncated gene clone to yield plasmid PAA352. The  
 CC clone was used to produce chimeric proteins by gene fusion with an  
 CC antigen coding sequence, e.g. the coding sequence of somatostatin,  
 CC gonadotropin releasing hormone or rotavirus viral protein 4, i.e.  
 CC leukotoxin works as a carrier protein to bring about a larger  
 CC immune response than the antigen alone. Immunisation with these  
 CC antigens can regulate growth rate, lactation and reproductive efficiency.  
 CC See also R34546-8.

SO Sequence 926 AA:

Query Match 59.2%; Score 2126; DB 14; Length 926; 1  
 Best Local Similarity 52.6%; Pred. No. 9,7e-135;  
 Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

OY 111 SPFKTAKKIILYIPONYDYDFEGNGLDLYKAAELGIEVORERNINATQTSLSLT 170
DB 8 sfptgakkillyipnygydeqngldlykaaeelgyevgreenmlataqtslgtl 67
OY 171 QTAIGTERGIYVSAPOIDKLLOKTAKGALGSAESIYONANKAKVLISGOSILASVLA 230
DB 68 qtaiglttergiylsapqidlklqktakgalaesiesivqanbakakvlisgilsvla 127
OY 231 GMDLDEALONNSQHALAKAGLELTNSLIENIANSVKTLDERGEQISOPGSKLONIKGLG 290
DB 128 gmdldealgnsqhalakagletnslileniansvktldetgeqisqfysklgnkgy 187
OY 291 TLGDKIKNIGGDKAGLGIDVISGLSGATAALVADKNASTAKKYGACFELANOVGNI 350
DB 188 tlgdkiknigldkaglgidvisgllsgataalvldknaastakkygagfelanvgvni 247
OY 351 TKAVSSYILAOVRAAGLSTGPAALIASTVSLATSPILFAGIAOKENAKSLSEYAEER 410
DB 248 tkavssyllaqrvaaglstcpvaaliasvslasplafagldkfnhakslesyaert 307
OY 411 KRLGYDGNMLAEYORGSTIDASVTALNTALAIAGVSAANA----- 454
DB 308 krlgydgnmlaeayrgtldasvralntalaalagvsaaagsviallvsylt 367
OY 455 ----- 454
DB 368 gyistllgyskgamfehvanklnkiveweknhgnyfengydarylanlqdmkflin 427
OY 455 ----- 454
DB 428 lnkeiaervialtqgqwdnigdlaglsrlgekvlsqkayvdafeeghklkadvld 487
OY 455 ----- 454
DB 488 sangliidvnsqkaktqhl1ftprlltpgtehrervgtqkyeyicklnnrdswwkltgd 547
OY 455 ----- 454
DB 548 aasstfdlnvvgrigieldnagnvktckekliaklggeddnvfvsgsteldggedyd 607
OY 455 ----- 454

```

Db 608 rvhysrgnygaltdatkebegsgyfvnrfvetgkalhevtsthalvgnreekleyrhs 667  
 QY 455 ----- 454  
 Db 668 mngnhagyytkdtlkaveeigtshndifkyskfndafnggdvdtldgndgndrlfgyk 727  
 QY 455 ----- 454  
 Db 728 gddlldgngddfidggkgndllhggkgddlfvhrkggdndiltdsgndklsfsgnuk 787  
 QY 455 DLTFEYKRNHLVTTNSKKEVTTIQNMFREADPKKEVPNKRATDEKIEELTIGNGERITS 514  
 Db 788 dlfevkvkhlvltnskkekvltqnmfiredafkevpnykatdekleeilgngerits 847  
 QY 515 KOYDDLIAGKNGKITODELSKVVDNELKHSKNVTNSLDKLSSVSAPFSSNDSSRNLV 574  
 Db 848 kyvddliakngkltqdeliskvvdnyellkhsknvnsldklssvsafssndssrnvlv 907  
 QY 575 APTSMLOSLSSLOFARGS 593  
 Db 908 aptsmldgsisslqfargs 926  
 RESULT 14  
 R50291  
 ID R50291 standard; Protein: 926 AA.  
 AC R50291;  
 DT 06-OCT-1994 (first entry)  
 DE Recombinant leukotoxin from plasmid PAA352.  
 XX  
 KW Vaccine; outer membrane protein; OMP; Haemophilus somnus;  
 KW Iron regulated protein; leukotoxin; Pasteurella haemolytica;  
 KM LKT352.  
 KM  
 XX  
 OS Pasteurella haemolytica A1 (strain B122).  
 XX  
 PN CA2099707-A.  
 PD 03-JAN-1994.  
 PF 29-JUN-1993; 93CA-2099707.  
 PR 02-JUL-1992; 92US-0908253.  
 PR  
 PA (UYSA-) UNITV SASKATCHEWAN.  
 XX  
 PI Harland RJ, Potter AA;  
 DR MPI: 1994-092909/12.  
 DR N-PSDB; Q44760.  
 XX  
 PT Haemophilus somnus outer membrane protein extract -  
 PT enriched with iron-regulated proteins, opt. contg.  
 PT Leuco:toxin antigens, for use as vaccine  
 XX  
 PS Claim 5; Fig 5; 78pp; English.  
 CC A vaccine comprising an outer membrane protein (OMP) extract of  
 CC Haemophilus somnus enriched with iron regulated proteins is new.  
 CC The vaccine pref. further comprises an immunogenic leukotoxin  
 CC polypeptide, esp. an immunogenic pasteurella haemolytica leukotoxin  
 CC homologous to LKT352. Example 1.2 describes the prodn. of  
 CC P. haemolytica recombinant leukotoxin from PAA352.  
 CC Two expression constructs were made. One, PAA342, contained the  
 CC 5'-Ahail fragment of the ltkA gene, while the other, PAA345,  
 CC contained the entire ltkA gene. PAA342 expressed a truncated  
 CC leukotoxin peptide at high levels, while PAA345 expressed full  
 CC length leukotoxin at very low levels. Therefore, the 3' end  
 CC of the ltkA gene was ligated into PAA342, yielding plasmid PAA352.  
 CC LKT352 or new leukotoxin is 98% homologous to authentic

CC leukotoxin.  
 CC NB: the protein sequence in Fig 5 comprises 926 amino acids,  
 CC however this protein is described in the text as having  
 CC 931 amino acids.  
 XX  
 SQ Sequence 926 AA:  
 Query Match 59.2%; Score 2126; DB 15; Length 926;  
 Best Local Similarity 52.6%; Pred. No. 9.7e-135;  
 Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;  
 QY 111 SFPKAGAKKILLYIPONYOYDEQNGLODLYKAAEELGIEVOREERNNTATAOTSIGTI 170  
 Db 8 sftkgakaklilylpnygydteqnglqdlvkaaeelgvevgreenmlatqtsigtl 67  
 QY 171 QTAIGLTERIVLSAQIDKLLOKTRAGQALSGABESIVQNAKAKVYLSIGTISVLA 230  
 Db 68 qtaiglttergvisapqidlqktkaggalgsaesivqnaakakvlsiglsivla 127  
 QY 231 GMDLDEALONSNQHLAKAGLELTNSLIENIANSVKTTDEPGEQISQPSKLONIKGLG 290  
 Db 128 gmdldealqnsnqhalakagletnsljeniansvktldeifeglsqfsgkignklyg 187  
 QY 291 TLGDKLKNIGGLDKAGLGLDVISGLSGATPAALVLDKNASTAKKYGACFELANQVGMT 350  
 Db 188 tlgdklknigglidkaglgldvisglsgataalvldknaastakkyvgaefelanvgnl 247  
 QY 351 TKAVSSYIIAQRYAAGLSSGPYAAALIASTVSLAISPLAFAGIADKFNNAKSLESTAEER 410  
 Db 248 tkavssyilaqrvaaglstgpyaaalastvslaisplafagiadkfnhakslesyaert 307  
 QY 411 KRLGYGDNLLAEYRGSTGTIDASYTAINTALAIAGVSAANA----- 454  
 Db 308 krlgygdnllaeqrgtltidasvtaintalaigvsaanaagvsapiallvsglt 367  
 QY 455 ----- 454  
 Db 368 gvtstllgyskgamfehvanklnhkiweweknhgknyfengydarylamlqdnmkflln 427  
 QY 455 ----- 454  
 Db 428 lnkelaervialtqgqwdnigdlaglsrlgekvlsqkayvdafeegkhikadklvqlid 487  
 QY 455 ----- 454  
 Db 488 sangliavnsqakqchllfrpplltptghehrervqtgkyeyitklhnrvdskitdg 547  
 QY 455 ----- 454  
 Db 548 aasftdltnvgigldelnagnvcktketkllaklgsgdnhvfvgsgteldggyvd 607  
 QY 455 ----- 454  
 Db 608 rvhysrgnygaltdatkebegsgyfvnrfvetgkalhevtsthalvgnreekleyrhs 667  
 QY 455 ----- 454  
 Db 668 mngnhagyytkdtlkaveeigtshndifkyskfndafnggdvdtldgndgndrlfgyk 727  
 QY 455 ----- 454  
 Db 728 gddlldgngddfidggkgndllhggkgddlfvhrkggdndiltdsgndklsfsgnuk 787  
 QY 455 DLTFEYKRNHLVTTNSKKEVTTIQNMFREADPKKEVPNKRATDEKIEELTIGNGERITS 514  
 Db 788 dlfevkvkhlvltnskkekvltqnmfiredafkevpnykatdekleeilgngerits 847  
 QY 515 KOYDDLIAGKNGKITODELSKVVDNELKHSKNVTNSLDKLSSVSAPFSSNDSSRNLV 574  
 Db 848 kyvddliakngkltqdeliskvvdnyellkhsknvnsldklssvsafssndssrnvlv 907  
 QY 575 APTSMLOSLSSLOFARGS 593

```

Db      908 aptsmldgslsifargs 926
|||||
RESULT 15
W03945
ID      W03945 standard; Protein: 926 AA.
XX
AC      W03945;
XX
DT      20-NOV-1996 (first entry)
XX
DE      P. haemolytica truncated leukotoxin (LKT352).
XX
KW      Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW      fusion protein; immunogen; vaccine; fertility control;
KW      contraceptive; sterilisation; plasmid pMA352.
XX
OS      Pasteurella haemolytica A1 strain B122.
XX
PN      W09624675-A1.
XX
PD      15-AUG-1996.
XX
PF      24-JAN-1996; 96MO-CA00049.
XX
PR      10-FEB-1995; 95US-0387156.
XX
PA      (UYSA-) UNIV SASKATCHEWAN.
XX
PI      Manns JG, Potter AA:
XX
DR      WPI: 1996-384447/38.
XX
N-PSDB: T37179.
XX
PT      Gonadotropin-releasing hormone multimer fusion proteins - with
PT      leukotoxin polypeptide for increased immunogenicity, useful in
XX      antifertility vaccine prodn.
XX
PS      Example 1; Fig 3A-3I; 87pp; English.
XX
CC      A truncated leukotoxin (W03945), LKT 352, lacks the cytotoxic
CC      portion of the native protein from Pasteurella haemolytica. It is
CC      the product of plasmid pMA352 which carries a truncated lktA gene
CC      (T37179). A fusion protein (W03942) between LKT352 and a
CC      gonadotropin releasing hormone tetramer can be expressed in
CC      Escherichia coli. This is useful as a vaccine for fertility
CC      control, partic. immunological sterilisation of domestic or
CC      farm animals.
XX
SQ      Sequence 926 AA:
Query Match 59.2%; Score 2126; DB 17; Length 926;
Best Local Similarity 52.6%; Pred. No. 9.7e-135;
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;
Db      111 SPPKGTAKKIIYIPONTQYDTREGNGIADLYKAABELGIEVOREBRNNIATPOTSIGTI 170
Db      8 sfpkcgakkiliypqnyqdydegngldlvkaaeelgievgreenmiataqtslgtl 67
QY      171 OTAIGTGERGIYLSAPOIDKLOKTKAGALGASISIVONANKAKTVLSGTSISVLA 230
Db      68 qtaiglttergiyvisapqldkllqkckagaglsaesilvqnankaktvlsqslgsvla 127
QY      231 GMDLDEALONNSQHALKAGLELTNSLTENTANSVKTLDERGEQISQFSKLQNIKGLG 290
Db      128 gmdldealgmsnqhalakagletnslientansvktldetgeqisqfsklqnlkglg 187
QY      291 TLGDKLKNIGGIDKAGLGIDVTSGLSGATAALYLADKNASTAKKVGAGFELANOVGNI 350
Db      188 tlgdklknigldkaglgldvisgllsgataalvladknastakkvgagfelanqvgnl 247

```

```

QY      351 TKAVSSYIIAORVAAGLSSTGPAAIIASTVSLAISPLAFACIADKFNHAKLSYAEERF 410
Db      248 tkavssyilaqrvaaglsstgpvaalaaivslaisplafagidckfmhaksleypf 307
QY      411 KRLGYDGNLAEYORGTTIDASVTAINTLAALAGVSAAA----- 454
Db      308 krlgydgnllaeYrgtctidasvtaintalaaiaagvsaaagvsiaaplallvsgit 367
QY      455 ----- 454
Db      368 gyistllgskqamfehvaanklnkiveweknhgknyfengydarylanlqdmkflln 427
QY      455 ----- 454
Db      428 lnkejaerviaitqgwdmniaglsrlgekvlsqkayvdafeegkhkacklvql 487
QY      455 ----- 454
Db      488 saanglidsnsgkaktqhlfrtplltpgtetrvrvqgkyeyicklnrvdswkldg 547
QY      455 ----- 454
Db      548 aasstfdltuvvgirigielhagavnlkethkilaigeddnvfvsgteldggyd 607
QY      455 ----- 454
Db      608 rvhysrnygaltdatketegsylvtrfvetgkalhevsthalvgnreekleyrhs 667
QY      455 ----- 454
Db      668 nqnhagytkdtkaveqilgtsnndlfkgsfndafngdvdvtdlgndndtlfggk 727
QY      455 ----- 454
Db      728 gddlldgngddfidgqkndllhggkgddlfvrhkgdndlitsdgnklsfsgnlnk 787
QY      455 DLTPFKVKHNLVTNTSKKEKVTIOMWPREADFAKVPYKATKDKREIITIGONGERTIS 514
Db      788 dltpfkvhnlvntskkrevlqmwpreadfakevpykakdekleeilsgnqnerlts 847
QY      515 KOVDLILKNGKITQDELskvvdnyELKHSKNTNSLDKLSSVSASFSSNDsrnlyv 574
Db      848 kqvddliakngkitqdelskvvdnyellhsknvtnsldklssvsafssndsrnlyv 907
QY      575 APTSMUDGSLSLQPARGS 593
Db      908 aptsmldgslsifargs 926

```

Search completed: March 2, 2001, 10:53:59  
Job time: 924 sec

**THIS PAGE BLANK (USPTO)**





```
Db 1 MATVIDRSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQ 60
QY 61 HWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSFPKTKAKKI 120
Db 61 HWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSFPKTKAKKI 120
QY 121 ILYIPONYOYDTBOGNLODLVKAABELGIEVOREERNINATATOTSLGTOTAIGLTERG 180
Db 121 ILYIPONYOYDTBOGNLODLVKAABELGIEVOREERNINATATOTSLGTOTAIGLTERG 180
QY 181 IYLSAPOIDKLOKTRAGQALGSAESIYONANKAKTVLSIGISILSVLAGMDLDEALON 240
Db 181 IYLSAPOIDKLOKTRAGQALGSAESIYONANKAKTVLSIGISILSVLAGMDLDEALON 240
QY 241 NSNOHALAKAGLELTNSLIENIANSVKTTIDEFGEQISQFSGSKLONIKGLTGLDKLKNIG 300
Db 241 NSNOHALAKAGLELTNSLIENIANSVKTTIDEFGEQISQFSGSKLONIKGLTGLDKLKNIG 300
QY 301 GLDKAGIGLDVIGSLSGATAALVLDKNASTAKKVGAGFELANOVVGNITTKAVSSYILA 360
Db 301 GLDKAGIGLDVIGSLSGATAALVLDKNASTAKKVGAGFELANOVVGNITTKAVSSYILA 360
QY 361 QRYAAGLSSTGPAALIASVSLAISPLAFAGIADKFENHAKSLESYAERPKKLGYGDNL 420
Db 361 QRYAAGLSSTGPAALIASVSLAISPLAFAGIADKFENHAKSLESYAERPKKLGYGDNL 420
QY 421 LAEYORGTGTIDASVTAINTALAIAGVSAAADLTFEEKVKNHLYITNSKKKEKVTIOMN 480
Db 421 LAEYORGTGTIDASVTAINTALAIAGVSAAADLTFEEKVKNHLYITNSKKKEKVTIOMN 480
QY 481 FRADPAKEVPNKATKDEKIEEIIICONGERTSKOYDLDLIAKNGKITODELSKYVDNY 540
Db 481 FRADPAKEVPNKATKDEKIEEIIICONGERTSKOYDLDLIAKNGKITODELSKYVDNY 540
QY 541 ELKHKSNTVNSLDKLISVSATFTSSNDSRNVLVAPTSMDOSSLISQFARGSQHWSYGL 600
Db 541 ELKHKSNTVNSLDKLISVSATFTSSNDSRNVLVAPTSMDOSSLISQFARGSQHWSYGL 600
QY 601 RPSGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMS 660
Db 601 RPSGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMS 660
QY 661 SYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSS 695
Db 661 SYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSS 695

RESULT 2
US-09-306-689-13
; Sequence 13, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; TITLE OF INVENTION: ANIMALS
; FILE REFERENCE: 9001-0047
; CURRENT APPLICATION NUMBER: US/09/306,689B
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/088,024
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: defines a leukotoxin GnRH chimera
US-09-306-689-13

Query Match 100.0%; Score 3591; DB 17; Length 695;
Best Local Similarity 100.0%; Pred. No. 3,le-281;
```

```
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATVIDRSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQ 60
Db 1 MATVIDRSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQ 60
QY 61 HWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSFPKTKAKKI 120
Db 61 HWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSFPKTKAKKI 120
QY 121 ILYIPONYOYDTBOGNLODLVKAABELGIEVOREERNINATATOTSLGTOTAIGLTERG 180
Db 121 ILYIPONYOYDTBOGNLODLVKAABELGIEVOREERNINATATOTSLGTOTAIGLTERG 180
QY 181 IYLSAPOIDKLOKTRAGQALGSAESIYONANKAKTVLSIGISILSVLAGMDLDEALON 240
Db 181 IYLSAPOIDKLOKTRAGQALGSAESIYONANKAKTVLSIGISILSVLAGMDLDEALON 240
QY 241 NSNOHALAKAGLELTNSLIENIANSVKTTIDEFGEQISQFSGSKLONIKGLTGLDKLKNIG 300
Db 241 NSNOHALAKAGLELTNSLIENIANSVKTTIDEFGEQISQFSGSKLONIKGLTGLDKLKNIG 300
QY 301 GLDKAGIGLDVIGSLSGATAALVLDKNASTAKKVGAGFELANOVVGNITTKAVSSYILA 360
Db 301 GLDKAGIGLDVIGSLSGATAALVLDKNASTAKKVGAGFELANOVVGNITTKAVSSYILA 360
QY 361 QRYAAGLSSTGPAALIASVSLAISPLAFAGIADKFENHAKSLESYAERPKKLGYGDNL 420
Db 361 QRYAAGLSSTGPAALIASVSLAISPLAFAGIADKFENHAKSLESYAERPKKLGYGDNL 420
QY 421 LAEYORGTGTIDASVTAINTALAIAGVSAAADLTFEEKVKNHLYITNSKKKEKVTIOMN 480
Db 421 LAEYORGTGTIDASVTAINTALAIAGVSAAADLTFEEKVKNHLYITNSKKKEKVTIOMN 480
QY 481 FRADPAKEVPNKATKDEKIEEIIICONGERTSKOYDLDLIAKNGKITODELSKYVDNY 540
Db 481 FRADPAKEVPNKATKDEKIEEIIICONGERTSKOYDLDLIAKNGKITODELSKYVDNY 540
QY 541 ELKHKSNTVNSLDKLISVSATFTSSNDSRNVLVAPTSMDOSSLISQFARGSQHWSYGL 600
Db 541 ELKHKSNTVNSLDKLISVSATFTSSNDSRNVLVAPTSMDOSSLISQFARGSQHWSYGL 600
QY 601 RPSGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMS 660
Db 601 RPSGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMS 660
QY 661 SYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSS 695
Db 661 SYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSS 695

RESULT 3
US-09-383-912-16
; Sequence 16, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentl Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-16

```

```

Query Match      100.0%; Score 3591; DB 17; Length 699;
Best Local Similarity 100.0%; Pred. No. 3.1e-281;
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MATVIDRSQHMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQ 60
DB 1 MATVIDRSQHMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQ 60
OY 61 HMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQ 120
DB 61 HMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQ 120
OY 121 ILIYPONYQYDTBOGNGIQLDYKAAEELGIVOREERNINATAOTSIGTITGALGERG 180
DB 121 ILIYPONYQYDTBOGNGIQLDYKAAEELGIVOREERNINATAOTSIGTITGALGERG 180
OY 181 IYLSAPOIDKLQTKKAGALSAESIYQNNKAKTVLSIGISILGSLVLAQMDDEALQN 240
DB 181 IYLSAPOIDKLQTKKAGALSAESIYQNNKAKTVLSIGISILGSLVLAQMDDEALQN 240
OY 241 NSNQHAKAGLELNSLIENIANSVKTLDFEGEIQSFGSKLQNKIGLGLGDKLNIG 300
DB 241 NSNQHAKAGLELNSLIENIANSVKTLDFEGEIQSFGSKLQNKIGLGLGDKLNIG 300
OY 301 GLDRAGLGLDVISGLSGATATLADKNASTAKKVGAGFELANOVGNITKAVSSYILA 360
DB 301 GLDRAGLGLDVISGLSGATATLADKNASTAKKVGAGFELANOVGNITKAVSSYILA 360
OY 361 ORVAAGLSSTGPVAALIASTVSLAISPFAAGIADKFNHAKSLESYARFKKLGIDGNL 420
DB 361 ORVAAGLSSTGPVAALIASTVSLAISPFAAGIADKFNHAKSLESYARFKKLGIDGNL 420
OY 421 LAEYQRTGTIDASTATINTALAAIAGVSAAADLTFEEKYHNLYITNSKKEKTYIONW 480
DB 421 LAEYQRTGTIDASTATINTALAAIAGVSAAADLTFEEKYHNLYITNSKKEKTYIONW 480
OY 481 FREDFAKEVNYKATKDEKIEEITIGNGERITSKQVDDLIAKNGKITODELSKYVDNY 540
DB 481 FREDFAKEVNYKATKDEKIEEITIGNGERITSKQVDDLIAKNGKITODELSKYVDNY 540
OY 541 ELLHRSKNTVSLDKLSSVSAFTSSNDSRVNVAPTSMLDOSLSSLOFARSGSHMSYGL 600
DB 541 ELLHRSKNTVSLDKLSSVSAFTSSNDSRVNVAPTSMLDOSLSSLOFARSGSHMSYGL 600
OY 601 RPSGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMS 660
DB 601 RPSGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMS 660
OY 661 SYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQ 695
DB 661 SYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQ 695

```

```

RESULT 4
US-09-383-912-10
; Sequence 10, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-10

```

```

Query Match      73.6%; Score 2642.5; DB 17; Length 544;
Best Local Similarity 99.4%; Pred. No. 9.4e-205;
Matches 534; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```

```

OY 111 SFPKTKAKKILLYIPONYQYDTBOGNGIQLDYKAAEELGIVOREERNINATAOTSIGT 170
DB 8 SFPKTKAKKILLYIPONYQYDTBOGNGIQLDYKAAEELGIVOREERNINATAOTSIGT 67
OY 171 QTAIGLTERGIVLSAPOIDKLQTKKAGALSAESIYQNNKAKTVLSIGISILGSLVLA 230
DB 68 QTAIGLTERGIVLSAPOIDKLQTKKAGALSAESIYQNNKAKTVLSIGISILGSLVLA 127
OY 231 GMDLDEALONNSNQHAKAGLELNSLIENIANSVKTLDFEGEIQSFGSKLQNKIGL 290
DB 128 GMDLDEALONNSNQHAKAGLELNSLIENIANSVKTLDFEGEIQSFGSKLQNKIGL 187
OY 291 TLGDKLNIGLGLDRAGLGLDVISGLSGATATLADKNASTAKKVGAGFELANOVGN 350
DB 188 TLGDKLNIGLGLDRAGLGLDVISGLSGATATLADKNASTAKKVGAGFELANOVGN 247
OY 351 TKAVSSYILAORVAAGLSSTGPVAALIASTVSLAISPFAAGIADKFNHAKSLESYARF 410
DB 248 TKAVSSYILAORVAAGLSSTGPVAALIASTVSLAISPFAAGIADKFNHAKSLESYARF 307
OY 411 KKLGDGNLLAEYQRTGTIDASTATINTALAAIAGVSAAAA---DLTFEEKYHNLYI 467
DB 308 KKLGDGNLLAEYQRTGTIDASTATINTALAAIAGVSAAAA---DLTFEEKYHNLYI 367
OY 468 TNSKKEKTYIONWTFREADFAKEVNYKATKDEKIEEITIGNGERITSKQVDDLIAKNGK 527

```

|||||  
Db 368 TNSKKEKVIQNNFREADFAKEVPNKATKDEKIEIEIIGONGERITSKYDDLIANGNGK 427  
QY 528 ITODELSKYVDNELLKHSKNVTNSLDKLSSVSAPFSSNDNRNVLYAPTSMDQSLSSL 567  
Db 428 ITODELSKYVDNELLKHSKNVTNSLDKLSSVSAPFSSNDNRNVLYAPTSMDQSLSSL 487  
QY 588 OFARQSQHMSYGLRPGSSQDMSYGLRPGSSQHMSYGLRPGSSQDMSYGLRPGGS 644  
Db 488 OFARQSQHMSYGLRPGSSQDMSYGLRPGSSQHMSYGLRPGSSQDMSYGLRPGGS 544  
RESULT 5  
US-09-383-912-8  
; Sequence 8, Application US/09383912  
; GENERAL INFORMATION:  
; APPLICANT: POTTER, ANDREW A.  
; APPLICANT: MANN, JOHN G.  
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: REED & ROBINS LLP  
; STREET: 285 HAMILTON AVENUE, SUITE 200  
; CITY: PALO ALTO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/383,912  
; FILING DATE:  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/694,865  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MCCracken, THOMAS P.  
; REGISTRATION NUMBER: 38,548  
; REFERENCE/DOCKET NUMBER: 9001-0016.22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)327-3400  
; TELEFAX: (415)327-3231  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 977 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-383-912-8  
Query Match 67.6%; Score 2426; DB 17; Length 977;  
Best Local Similarity 55.1%; Pred. No. 7.8e-187;  
Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

Db 188 TIGDRLKNIIGLDKAGLDVIGSLGATPAALVLADKNASTAKVAGFELANOVVNT 247  
QY 351 TKAVSSYTLAQRVAGLSSGPAALIASTVSLATISPLAFAGIADKFNHAKSLESYAE 410  
Db 248 TKAVSSYTLAQRVAGLSSGPAALIASTVSLATISPLAFAGIADKFNHAKSLESYAE 307  
QY 411 KKLGDNDMLAEYORGTGTIDASVTAINPTALAIAGVSAAAA 454  
Db 308 KKLGDNDMLAEYORGTGTIDASVTAINPTALAIAGVSAAAA 367  
QY 455 454  
Db 368 GVISTILLQYSQAMFEHVANKLNKIVEKNNHKNFEENGYDARYLANLDNMKFLN 427  
QY 455 454  
Db 428 LNKELDAERYIAITQQQMDNNIGDLAGISRLGKVLGSKAYVAFEGEKHAKDLVOLD 487  
QY 455 454  
Db 488 SANGIIDVNSGKAKTQHILFRTPLLPGETHEHREVOTGKEYITKLNINRVDNWKITDG 547  
QY 455 454  
Db 548 AASTFDLTNVVORIGIELDNAGNVTKETKLIKALGEGDNNVFGSGTTEIDGEGYD 607  
QY 455 454  
Db 608 RVHYSKNGNVALTIDATKTEGQSTYVNRVETGKALHEVTSHTHTALVGNREKIEYRHS 667  
QY 455 454  
Db 668 NNGHAGYYTKDTLKAVEELIGTSHNDIFPKSGKFNDAFNGCGDVTIDGNDRLFGGK 727  
QY 455 454  
Db 728 GDDILDGNGDDFIDGKGNLDLHGKGDDIFVHRKGDNDITDSDGNDKLSFSDSNK 787  
QY 455 DLTFEYKHNLTNTSKKKEKVTIONMFREADFAKEVPNKATKDEKIEIEIIGONGERITS 514  
Db 788 DLTFEYKHNLTNTSKKKEKVTIONMFREADFAKEVPNKATKDEKIEIEIIGONGERITS 847  
QY 515 KQYDDLIANGNGKITODELSKYVDNELLKHSKNVTNSLDKLSSVSAPFSSNDNRNVLY 574  
Db 848 KQYDDLIANGNGKITODELSKYVDNELLKHSKNVTNSLDKLSSVSAPFSSNDNRNVLY 907  
QY 575 APTSMIDQSLSSLOFARQSQHMSYGLRPGSSQDMSYGLRPGSSQHMSYGLRPGGSQD 634  
Db 908 APTSMIDQSLSSLOFARQSQHMSYGLRPGSSQDMSYGLRPGSSQHMSYGLRPGGSQD 967  
QY 635 WSYGLRPGGS 644  
Db 968 WSYGLRPGGS 977  
RESULT 6  
US-09-252-149A-26  
; Sequence 26, Application US/09252149A  
; GENERAL INFORMATION:  
; APPLICANT: Barker, Christopher A.  
; APPLICANT: Morsey, Mohamed  
; TITLE OF INVENTION: IMMUNOLOGICAL METHODS TO MODULAR MYOSTATIN IN  
; TITLE OF INVENTION: VERTEBRATE SUBJECTS  
; FILE REFERENCE: 9001-0042  
; CURRENT APPLICATION NUMBER: US/09/252,149A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: 60/075,213  
; PRIOR FILING DATE: 1998-02-19  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 490  
; TYPE: PRT

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: leukotoxin polypeptide carrier  
US-09-252-149A-26

Query Match 65.6%; Score 2354; DB 16; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1,7e-181;  
Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 SPPKTKAKKIIILYIPQNYQYDTEQNGLODLYKAAEELGIEVQREERNINATAQTSLSGT 170  
DB 8 SPPKTKAKKIIILYIPQNYQYDTEQNGLODLYKAAEELGIEVQREERNINATAQTSLSGT 67  
QY 171 QTAIGLTERGIVLSAPQIDKLOKTKRAGALGSAESIYONANKAKTVLSGIOSILGSVLA 230  
DB 68 QTAIGLTERGIVLSAPQIDKLOKTKRAGALGSAESIYONANKAKTVLSGIOSILGSVLA 127  
QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENTANSVKTLDEFGEOISQFGSKLQNIKGLG 290  
DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENTANSVKTLDEFGEOISQFGSKLQNIKGLG 187  
QY 291 TLGDKLKNIGGLDKAGLGIDVYSGILSGATATVLAADKNASTAKKVGAGFELANQVGN 350  
DB 188 TLGDKLKNIGGLDKAGLGIDVYSGILSGATATVLAADKNASTAKKVGAGFELANQVGN 247  
QY 351 TKAVSSYLLAQRVAAGLSSTGPAALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 410  
DB 248 TKAVSSYLLAQRVAAGLSSTGPAALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 307  
QY 411 KRLGYDGNLLAEYQRGCTIDASTATNTALAAIAGVSAADLTPEKYKHNLYTNS 470  
DB 308 KRLGYDGNLLAEYQRGCTIDASTATNTALAAIAGVSAADLTPEKYKHNLYTNS 367  
QY 471 KKEKTYIONMFREADFAEVPYKATKDEKIEIIGONGERTSKOVODLLAKNGKITQ 530  
DB 368 KKEKTYIONMFREADFAEVPYKATKDEKIEIIGONGERTSKOVODLLAKNGKITQ 427  
QY 531 DELSKRVNDYELLKHSKNVTNSLDKLSSVSFTSSNDSRNVLVAPTSMLDQSLSSLOFA 590  
DB 428 DELSKRVNDYELLKHSKNVTNSLDKLSSVSFTSSNDSRNVLVAPTSMLDQSLSSLOFA 487  
QY 591 RGS 593  
DB 488 RGS 490

RESULT 7  
US-09-252-149B-26  
Sequence 26, Application US/09252149B  
GENERAL INFORMATION:  
APPLICANT: Barker, Christopher A.  
APPLICANT: Morsey, Mohamad  
TITLE OF INVENTION: IMMUNOLOGICAL METHODS TO MODULAR MYOSTATIN IN  
TITLE OF INVENTION: VERTEBRATE SUBJECTS  
FILE REFERENCE: 9001-0042  
CURRENT APPLICATION NUMBER: US/09/252,149B  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: 60/075,213  
PRIOR FILING DATE: 1998-02-19  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 26  
LENGTH: 490  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: leukotoxin polypeptide carrier  
US-09-252-149B-26

Query Match 65.6%; Score 2354; DB 27; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1,7e-181;  
Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 SPPKTKAKKIIILYIPQNYQYDTEQNGLODLYKAAEELGIEVQREERNINATAQTSLSGT 170  
DB 8 SPPKTKAKKIIILYIPQNYQYDTEQNGLODLYKAAEELGIEVQREERNINATAQTSLSGT 67  
QY 171 QTAIGLTERGIVLSAPQIDKLOKTKRAGALGSAESIYONANKAKTVLSGIOSILGSVLA 230  
DB 68 QTAIGLTERGIVLSAPQIDKLOKTKRAGALGSAESIYONANKAKTVLSGIOSILGSVLA 127  
QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENTANSVKTLDEFGEOISQFGSKLQNIKGLG 290  
DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENTANSVKTLDEFGEOISQFGSKLQNIKGLG 187  
QY 291 TLGDKLKNIGGLDKAGLGIDVYSGILSGATATVLAADKNASTAKKVGAGFELANQVGN 350  
DB 188 TLGDKLKNIGGLDKAGLGIDVYSGILSGATATVLAADKNASTAKKVGAGFELANQVGN 247  
QY 351 TKAVSSYLLAQRVAAGLSSTGPAALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 410  
DB 248 TKAVSSYLLAQRVAAGLSSTGPAALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 307  
QY 411 KRLGYDGNLLAEYQRGCTIDASTATNTALAAIAGVSAADLTPEKYKHNLYTNS 470  
DB 308 KRLGYDGNLLAEYQRGCTIDASTATNTALAAIAGVSAADLTPEKYKHNLYTNS 367  
QY 471 KKEKTYIONMFREADFAEVPYKATKDEKIEIIGONGERTSKOVODLLAKNGKITQ 530  
DB 368 KKEKTYIONMFREADFAEVPYKATKDEKIEIIGONGERTSKOVODLLAKNGKITQ 427  
QY 531 DELSKRVNDYELLKHSKNVTNSLDKLSSVSFTSSNDSRNVLVAPTSMLDQSLSSLOFA 590  
DB 428 DELSKRVNDYELLKHSKNVTNSLDKLSSVSFTSSNDSRNVLVAPTSMLDQSLSSLOFA 487  
QY 591 RGS 593  
DB 488 RGS 490

RESULT 8  
US-07-779-171-10  
Sequence 10, Application US/07779171  
GENERAL INFORMATION:  
APPLICANT: Potter, Andrew A.  
APPLICANT: Redmond, Mark J.  
APPLICANT: Hughes, Huw P. A.  
TITLE OF INVENTION: ENHANCED IMMUNOCENICITY USING  
TITLE OF INVENTION: PASTEURELLA HEAMOLYTICA LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 545 Middlefield Road, suite 200  
CITY: Menlo Park  
STATE: California  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/779,171  
FILING DATE: 19911016  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Robins, Roberta L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 29310-2001600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-327-7250

```

: TELEFAX: 415-327-2951
:
: TELE: 706141
:
: INFORMATION FOR SEQ ID NO: 10:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 936 amino acids
:
: TYPE: AMINO ACID
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: US-07-779-171-10

```

Query Match	61.0%;	Score 2189;	DB 3;	Length 936;
Best Local Similarity	53.1%;	Pred. No. 1e-167;		
Matches 493; Conservative	0;	Mismatches 0;	Indels 436;	Gaps 1

[illegible]

```

Db      848  KQVDDLLAANGKNGKITQDELSKVYVNIELIKHSKNVTNSLDKLSSVSAFTSSNDNRNV  908
QY      575  APTSMLDGSLSLQFARGSQHWSTGLRPG  603
      |||||
Db      908  APTSMLDGSLSLQFARGSQHWSTGLRPG  936

```

RESULT 9  
US-08-455-970-12  
; Sequence 12, Application US/08455970

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/06/455,970  
 FILING DATE: 31-MAY-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/960,932  
 FILING DATE: 14-OCT-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ROHNS, ROBERTA L.  
 REGISTRATION NUMBER: 33,208  
 REFERENCE/DOCKET NUMBER: 9001-0016.10  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 327-3400  
 TELEFAX: (415) 327-3231  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 936 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-455-970-12

Query Match	61.0%;	Score 2189;	DB 8;	Length 936;
Best Local Similarity	53.1%;	Pred. No. 1e-167;		
Matches 493;	Conservative	0;	Mismatches 0;	Indels 436;
				Gaps 1

Oy	111	SEPKTGAKKILLYIPONQYOTDEGNGJODLYKAAEELGJEVOEENEENNTATQTSIGTI	170
Dd	8	SFPKGTGAKKILLYIPONTQYOTDEGNGJODLYKAAEELGJEVOEENEENNTATQTSIGTI	67
Oy	171	QTAIGLTERGIVLSAPOIDKLLQTKKAGQALSAESIIVONANKAKTVLSGIOSILGSVLA	230
Dd	68	QTAIGLTERGIVLSAPOIDKLLQTKKAGQALSAESIIVONANKAKTVLSGIOSILGSVLA	127
Oy	231	GMDDEALQNNNSNOHALAKAGLELTNSLIENIANSVYTTDFEFGQISQPSKXIONTKGIG	230
Dd	128	GMDDEALQNNNSNOHALAKAGLELTNSLIENIANSVYTTDFEFGQISOFPKSKIONTKGIG	187
Oy	291	TUGDKLNKIGGLDRAKGLDVIYSGLISGATATVLAJADKNASTAKKVGAGELANQYVYNI	350
Dd	188	TUGDKLNKIGGLDRAKGLDVIYSGLISGATATVLAJADKNASTAKKVGAGELANQYVYNI	247
Oy	351	TKAVSSYTLAORVAGLSSGTGPVALLIASTVGLAISPLRAGIADKRENNAKSLIESYAEERF	410

```

Db 248 TRAVSSYLLAORVAAGLSTSGVAAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF
OY 411 KKLGVGDGDLLEAYORGTIDASTATATATATAIAGVSAASA-----
Db 308 KKLGVGDGDLLEAYORGTIDASTATATATAIAGVSAASAASVIAPIALLVSGIT
OY 455 -----
Db 368 GVIITLLQYKQAMFEHVANKIHKIVEMKNNHGKNTFENGIDARYLANIADNMKFLIN
OY 455 -----
Db 428 LNKELQAEERVAITQOQWNNINIGLAGISRLGEKVLGSKAYDAFECHKIKADKLYOLD
OY 455 -----
Db 488 SANGIIDVNSGKAKTOHILFRTPLTPGTEHREVRQYQKYEYITKLNINRVDWMTDGG
OY 455 -----
Db 548 AASTFEDLTNVQRIEILDNAGNVTKTETKIATKLGEGDNNVFGSGTTEIDGEGYD
OY 455 -----
Db 608 RVHYSRGNYGALTIDATKETEGSYTVNRFVETGKALHEVTSHTALVGNREKIEYRHS
OY 455 -----
Db 668 NNQHHAGYTTKDTLKAVEEIIIGTSHNDIFKSGKFNDAFNGDGVDTIDGNDRLFGGK
OY 455 -----
Db 728 GDDIILDGNGDDFDIGGKGNLDLHGKGDGDI FVHRKGGDNDIITDSGNDKLSFSDSNLK
OY 455 -----
Db 788 DLFERKVNHNLYTNSKKEKVTIOMMFREADFAKEVPYNTKATKDEKIEIIGONGERITS
OY 455 -----
Db 847 DLFERKVNHNLYTNSKKEKVTIOMMFREADFAKEVPYNTKATKDEKIEIIGONGERITS
OY 515 KOVDLLIAGKNGKITODELSKYVDNVELLKHKNVNSIDKLISVSASFSSNDSRNVLY
Db 848 KOVDLLIAGKNGKITODELSKYVDNVELLKHKNVNSIDKLISVSASFSSNDSRNVLY
OY 575 APTSMIDQSLSLQFARGSOHWSYGLRPG 603
Db 908 APTSMIDQSLSLQFARGSOHWSYGLRPG 936

```

```

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 29310-2001600
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 926 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-779-171-14

Query Match 59.2%; Score 2126; DB 3; Length 926;
Best Local Similarity 52.6%; Pred. No. 1.3e-162;
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

OY 111 SEPKTGARKIILYIPQNYQYDTEQNGGLQDLYKAAEEIGIEVQREERNINATAQTSIGTI
Db 8 SEPKTGARKIILYIPQNYQYDTEQNGGLQDLYKAAEEIGIEVQREERNINATAQTSIGTI
OY 171 QTAIGLTERGIVLSAPQIDKLLQTKAGQALGSASIVQNNKAKTVLSGIQSLGSVLA
Db 68 QTAIGLTERGIVLSAPQIDKLLQTKAGQALGSASIVQNNKAKTVLSGIQSLGSVLA
OY 231 GMDIDELALONNSNOHALKAGLELTNSLIENIANSVKTLDFEGQISQFSGKLQNIKGLG
Db 128 GMDIDELALONNSNOHALKAGLELTNSLIENIANSVKTLDFEGQISQFSGKLQNIKGLG
OY 291 TLGDKLKNIGGLDKAGLGLDVIISGLSGATAALVLDKRNASTAKKVGAFELANQVGN
Db 188 TLGDKLKNIGGLDKAGLGLDVIISGLSGATAALVLDKRNASTAKKVGAFELANQVGN
OY 351 TKAVSSYLLAORVAAGLSTSGVAAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF
Db 248 TKAVSSYLLAORVAAGLSTSGVAAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF
OY 411 KKLGVGDGDLLEAYORGTIDASTATATATAIAGVSAASA-----
Db 308 KKLGVGDGDLLEAYORGTIDASTATATATAIAGVSAASAASVIAPIALLVSGIT
OY 455 -----
Db 368 GVIITLLQYKQAMFEHVANKIHKIVEMKNNHGKNTFENGIDARYLANIADNMKFLIN
OY 455 -----
Db 428 LNKELQAEERVAITQOQWNNINIGLAGISRLGEKVLGSKAYDAFECHKIKADKLYOLD
OY 455 -----
Db 488 SANGIIDVNSGKAKTOHILFRTPLTPGTEHREVRQYQKYEYITKLNINRVDWMTDGG
OY 455 -----
Db 548 AASTFEDLTNVQRIEILDNAGNVTKTETKIATKLGEGDNNVFGSGTTEIDGEGYD
OY 455 -----
Db 608 RVHYSRGNYGALTIDATKETEGSYTVNRFVETGKALHEVTSHTALVGNREKIEYRHS
OY 455 -----
Db 668 NNQHHAGYTTKDTLKAVEEIIIGTSHNDIFKSGKFNDAFNGDGVDTIDGNDRLFGGK
OY 455 -----
Db 728 GDDIILDGNGDDFDIGGKGNLDLHGKGDGDI FVHRKGGDNDIITDSGNDKLSFSDSNLK

```

```

RESULT 10
US-07-779-171-14
: Sequence 14, Application US/0779171
: GENERAL INFORMATION:
: APPLICANT: Potter, Andrew A.
: APPLICANT: Redmond, Mark J.
: APPLICANT: Hughes, Huw P. A.
: TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morrison & Foerster
: STREET: 545 Middlefield Road, Suite 200
: CITY: Menlo Park
: STATE: California
: COUNTRY: USA
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/779,171
: FILING DATE: 19911016

```

QY 455 DLFEEVKHNLVTNSKKEKVTIONMFREADFAKEVPNYKATDEKIEEIIIGNGERITS 514  
|||||  
Db 788 DLFEEVKHNLVTNSKKEKVTIONMFREADFAKEVPNYKATDEKIEEIIIGNGERITS 847  
QY 515 KOYDDLIAKNGKITODELSKYVDNELKHSKNVTNSLDKLISSVSAFTSSNDSRNVLV 574  
|||||  
Db 848 KOYDDLIAKNGKITODELSKYVDNELKHSKNVTNSLDKLISSVSAFTSSNDSRNVLV 907  
QY 575 APTSMLOSLSLOFARGS 593  
|||||  
Db 908 APTSMLOSLSLOFARGS 926  
RESULT 11  
US-08-455-970-2  
Sequence 2, Application US/08455970  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN  
TITLE OF INVENTION: CHIMERAS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,970  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 424  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 327-3400  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 926 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-455-970-2  
Query Match 59.2%; Score 2126; DB 8; Length 926;  
Best Local Similarity 52.6%; Pred. No. 1,3e-162;  
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;  
QY 111 SPFKTAKKILITIPONYQYDTFOGNGLODLYKAAEELGIEVOREPNNNTATQTSIGTI 170  
|||||  
Db 8 SPFKTAKKILITIPONYQYDTFOGNGLODLYKAAEELGIEVOREPNNNTATQTSIGTI 67  
QY 171 QTAIGLTERGIVLSAPOIDKLOKTRAGQALGSAESIYQANAKAKTVLSIGISLSVLA 230  
|||||  
Db 68 QTAIGLTERGIVLSAPOIDKLOKTRAGQALGSAESIYQANAKAKTVLSIGISLSVLA 127  
QY 231 GMDLDELALONNSQOHALAKAGLELNSLIENIANSVKTTLDFEGEOISQFGSKLQNTKGLG 290  
|||||

Db 128 GMDLDELALONNSQOHALAKAGLELNSLIENIANSVKTTLDFEGEOISQFGSKLQNTKGLG 187  
QY 291 TLGDKLKNIGGLDAGLGLDVISGLSGATPAALVLAADKNASTAKKVGAGFELANQVGN 350  
|||||  
Db 188 TLGDKLKNIGGLDAGLGLDVISGLSGATPAALVLAADKNASTAKKVGAGFELANQVGN 247  
QY 351 TKAVSSYIIAQRVAAGLSSTGPVAAIIASTVSLAISPLAFIADKFNHAKSLESYAE 410  
|||||  
Db 248 TKAVSSYIIAQRVAAGLSSTGPVAAIIASTVSLAISPLAFIADKFNHAKSLESYAE 307  
QY 411 KRIGYDNLATYORGCTGIDASVTAINMTALAIAGVSAANA 454  
|||||  
Db 308 KRIGYDNLATYORGCTGIDASVTAINMTALAIAGVSAANA 367  
QY 455 455 454  
Db 368 GVISTILLOYSKQAMFEHVANKINKIVEMKNNHGNFENGIDARYLANLQDNMFLN 427  
QY 455 455 454  
Db 428 LNKELQAEVYIATTOQWQNDNIGDLAIGSRLEKVLGKAYVDAFEGKHAKDLVOLD 487  
QY 455 455 454  
Db 488 SANGIIDVNSGKAKTOHILFRTPLTPGTEREVRVOTGKEYITRLNINRVDWKRITDG 547  
QY 455 455 454  
Db 548 AASTFDLIRNVORIGIELDNAGNVTIKETKIIAALGEGDDNVFVSGSTTEIDGEGVD 607  
QY 455 455 454  
Db 608 RVHSGNGALITIDATKETEQGSYTVNRFVETGKALHEVTSHTALVGNREKIEYRHS 667  
QY 455 455 454  
Db 668 NNQHAQYTKDTLKAVEELIGTSHNDIFKSGKENDAFNGGVDVTIDGNDGNDRLFGK 727  
QY 455 455 454  
Db 728 GDIIDGNGDDPFDGKGNLDLHGKGGDIIVHRKGGNDITDSDGNDKLSFSDSNLK 787  
QY 455 DLFEEVKHNLVTNSKKEKVTIONMFREADFAKEVPNYKATDEKIEEIIIGNGERITS 514  
|||||  
Db 788 DLFEEVKHNLVTNSKKEKVTIONMFREADFAKEVPNYKATDEKIEEIIIGNGERITS 847  
QY 515 KOYDDLIAKNGKITODELSKYVDNELKHSKNVTNSLDKLISSVSAFTSSNDSRNVLV 574  
|||||  
Db 848 KOYDDLIAKNGKITODELSKYVDNELKHSKNVTNSLDKLISSVSAFTSSNDSRNVLV 907  
QY 575 APTSMLOSLSLOFARGS 593  
|||||  
Db 908 APTSMLOSLSLOFARGS 926  
RESULT 12  
US-09-383-912-6  
Sequence 6, Application US/09383912  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/383,912  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 926 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-383-912-6

Query Match 59.2%; Score 2126; DB 17; Length 926;

Best Local Similarity 52.6%; Pred. No. 1.3e-162; Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

Db 111 SPPKTKAKIILYIPONTQYDTEOGNGIJDLYKAAEELGIEVOREERNINATAOTSIGTI 170  
| 111 SPPKTKAKIILYIPONTQYDTEOGNGIJDLYKAAEELGIEVOREERNINATAOTSIGTI 170  
| 8 SPPKTKAKIILYIPONTQYDTEOGNGIJDLYKAAEELGIEVOREERNINATAOTSIGTI 67  
Qy 171 QTAIGLTERGIVLSAPQIDKLTQKTKAQAALGSAESIYQANNAKKTIVLSGIOSILGSVLA 230  
| 171 QTAIGLTERGIVLSAPQIDKLTQKTKAQAALGSAESIYQANNAKKTIVLSGIOSILGSVLA 230  
| 68 QTAIGLTERGIVLSAPQIDKLTQKTKAQAALGSAESIYQANNAKKTIVLSGIOSILGSVLA 127  
Db 231 GMDIDEALONNSNOHALKAGLELTNSLIENANSVKTLDGEGEIOSFGSKLQNIKIG 290  
| 128 GMDIDEALONNSNOHALKAGLELTNSLIENANSVKTLDGEGEIOSFGSKLQNIKIG 187  
Qy 291 TLGDKLTNIGIDKAGLGLDIVISGLSGATAALVADKNASTAKVAGAFELANOVGNI 350  
| 188 TLGDKLTNIGIDKAGLGLDIVISGLSGATAALVADKNASTAKVAGAFELANOVGNI 247  
Qy 351 TKAVSYTLLAORVAAGLSSTGPAALINASTVSLAISPAPAGIADKFNHANSLESYAEERF 410  
| 248 TKAVSYTLLAORVAAGLSSTGPAALINASTVSLAISPAPAGIADKFNHANSLESYAEERF 307  
Qy 411 KKLGYDGNLLAEYORGTGTIDASTAINTALAIAGVSAASAAA----- 454  
| 308 KKLGYDGNLLAEYORGTGTIDASTAINTALAIAGVSAASAAA----- 367  
Qy 455 ----- 454  
Db 368 GVSTIILQSKQAMEHVAANKIHKIVEMEKNNHKNYFENGCDARYLANODNNKFFLLN 427  
| 455 ----- 454  
Qy 428 LNKELAERVAITQOQMDNINIGLAGISRLGEKVLGKAVVDAFEESGHIKADKLVDLD 487  
| 455 ----- 454  
Db 488 SANGIIVDSNGKAKTOHILFRTPLLPTEGHERERVOTGKEYITKININRVDSKTTDG 547  
| 455 ----- 454  
Qy 548 AASSTFDLTNVQRIQIELDNAGNVTKTETKRIIAKLGEQDNVYVSGTTEIDGCEGYD 607  
| 455 ----- 454  
Db 608 RVHYSRGNYGALLTDATKETEGSYTVNRPVETGKALHEVTSHTALVGNREKIEYRHS 667

Qy 455 ----- 454  
Db 668 NNOHAGYTTKDTLAAVEIITGTSNDIFKSGKFNDARNGGDVDTIGDNGDNRLFGCK 727  
| 455 ----- 454  
Qy 728 GDDIILDGNGDGFIDGKGNLHLHGKGGDDIFVHRKGGNDIITDSDGNDKLSFSDSNLK 787  
Qy 455 DTFEKKVHNLYITNSKKEKVTIQWFRPADFAKVPYKATKDEKIEIIGQNGERITS 514  
| 788 DTFEKKVHNLYITNSKKEKVTIQWFRPADFAKVPYKATKDEKIEIIGQNGERITS 847  
Qy 515 KOVDLLKNGKTIQDELSKVVDNYEILKSKNTNSLDKLISSVSAPTSNDSRNLY 574  
| 848 KOVDLLKNGKTIQDELSKVVDNYEILKSKNTNSLDKLISSVSAPTSNDSRNLY 907  
Qy 575 APTSMIDQSLSSLOFARGS 593  
| 908 APTSMIDQSLSSLOFARGS 926  
Db

## RESULT 13

US-07-779-171-12  
Sequence 12, Application US/07779171

## GENERAL INFORMATION:

APPLICANT: Potter, Andrew A.  
APPLICANT: Redmond, Mark J.  
APPLICANT: Hughes, Huw P. A.  
TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING  
TITLE OF INVENTION: PASTEURELLA HEMOLYTICA LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Morrison & Foerster  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: California  
COUNTRY: USA  
ZIP: 94025

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/779,171  
FILING DATE: 19911016  
CLASSIFICATION: 424

## ATTORNEY/AGENT INFORMATION:

NAME: Robins, Roberta L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 29310-2001600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-327-7250  
TELEFAX: 415-327-2951  
TELEX: 706141

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 943 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-779-171-12

Query Match 59.2%; Score 2126; DB 3; Length 943;

Best Local Similarity 52.6%; Pred. No. 1.3e-162; Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

Qy 111 SPPKTKAKIILYIPONTQYDTEOGNGIJDLYKAAEELGIEVOREERNINATAOTSIGTI 170  
| 111 SPPKTKAKIILYIPONTQYDTEOGNGIJDLYKAAEELGIEVOREERNINATAOTSIGTI 170  
| 8 SPPKTKAKIILYIPONTQYDTEOGNGIJDLYKAAEELGIEVOREERNINATAOTSIGTI 67  
Db 171 QTAIGLTERGIVLSAPQIDKLTQKTKAQAALGSAESIYQANNAKKTIVLSGIOSILGSVLA 230

```

|||||
DB 68 QTAIGLTERIVISAPOIDKLLOKTRAGQALGSAESIYQANAKAKTVLSGOSILSVLA 127
231 GMDLDELQNNNOHALAKAGLELFTNSLIENINSVKTDEFEQOISQFSKLONIKGLG 290
DB 128 GMDLDELQNNNOHALAKAGLELFTNSLIENINSVKTDEFEQOISQFSKLONIKGLG 187
291 TLGDKLKNIGGLDKAGIGLDIVISGLSGATPAALVLADKNASTAKKYGAGFELANOVVNT 350
DB 188 TLGDKLKNIGGLDKAGIGLDIVISGLSGATPAALVLADKNASTAKKYGAGFELANOVVNT 247
351 TKAVSSYIIAQRVAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAEERF 410
DB 248 TKAVSSYIIAQRVAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAEERF 307
411 KKLGYDGNLLAEYORGTGTIDASVTAINTALAAIGVSAASAAA----- 454
DB 308 KKLGYDGNLLAEYORGTGTIDASVTAINTALAAIGVSAASAAAGSVIASPALLVSGIT 367
455 ----- 454
DB 368 GVISTILQYSKOAMFEHVANKIHNKIVEMEKNNHKNYFENGYDARYLANIQQNMKFLN 427
455 ----- 454
DB 428 LNKELQERVIATQQQWNNIGDLAGISRLGKRVLSKRAYVDAFEBGKHAKDLVOLD 487
455 ----- 454
DB 488 SANGIIVDSNSGAKTOHILFRPLPLPGTEHREVRQTKYEYITTKLINRVDSWKITDG 547
455 ----- 454
DB 548 AASTFDLTINVQRIEELDNAGNVTKEKTKLIAKLGEEDDNVFEVSGTTEIDGGRYD 607
455 ----- 454
DB 608 RVHYSKNGALTIDATKETEQGSYTVNREVTGKALHEVTSHTALVGNREKIEYRHS 667
455 ----- 454
DB 668 NNQHAQYITKTLKAVEELITGSHNDIFKGSFNDAPNGDGVDTIDGNDGDLFGCK 727
455 ----- 454
DB 728 GDDILDGNGDDFLDGKGNLDLHGKGDIDFVHRKGDNDITDSDGNDKLSFSDSNLK 787
455 ----- 454
DB 455 DLFEKYNHMLVITNSKKEVYTIQNFREADPAKEVNTKATKDEKIEELIIGONGERITS 514
788 DLFEKYNHMLVITNSKKEVYTIQNFREADPAKEVNTKATKDEKIEELIIGONGERITS 847
515 KOYDDILAKNGKITODELSKYVDNVELLHKSNNVTNSLDKLISVSASFSSNDSRVLY 574
DB 848 KOYDDILAKNGKITODELSKYVDNVELLHKSNNVTNSLDKLISVSASFSSNDSRVLY 907
575 APTSMLDQSLSSLOFARGS 593
DB 908 APTSMLDQSLSSLOFARGS 926

```

RESULT 14  
US-08-455-970-10  
Sequence 10, Application US/08455970

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: REDMOND, MARK J.

APPLICANT: HUGHES, HOW P.A.

TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: REED & ROBINS

STREET: 285 HAMILTON AVENUE, SUITE 200

```

? CITY: PALO ALTO
? STATE: CALIFORNIA
? COUNTRY: UNITED STATES OF AMERICA
? ZIP: 94301
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/455,970
? FILING DATE: 31-MAY-1995
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 424
? FILING DATE: 14-OCT-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: ROBINS, ROBERTA L.
? REGISTRATION NUMBER: 33,208
? REFERENCE/DOCKET NUMBER: 9001-0016.10
? TELEPHONE: (415) 327-3400
? TELEFAX: (415) 327-3231
? INFORMATION FOR SEQ. ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 943 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-455-970-10

```

Query Match 59.2%; Score 2126; DB 8; Length 943;  
Best local similarity 52.6%; Pred. No. 1.3e-162;  
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

OY 111 SPPKGAKKIILYIPONYOYDTEQNGSLDVLKRAEELGTEVOREERNNTATQTSIGTI 170
DB 8 SPPKGAKKIILYIPONYOYDTEQNGSLDVLKRAEELGTEVOREERNNTATQTSIGTI 67
OY 171 QTAIGLTERIVISAPOIDKLLOKTRAGQALGSAESIYQANAKAKTVLSGOSILSVLA 230
DB 68 QTAIGLTERIVISAPOIDKLLOKTRAGQALGSAESIYQANAKAKTVLSGOSILSVLA 127
OY 231 GMDLDELQNNNOHALAKAGLELFTNSLIENINSVKTDEFEQOISQFSKLONIKGLG 290
DB 128 GMDLDELQNNNOHALAKAGLELFTNSLIENINSVKTDEFEQOISQFSKLONIKGLG 187
OY 291 TLGDKLKNIGGLDKAGIGLDIVISGLSGATPAALVLADKNASTAKKYGAGFELANOVVNT 350
DB 188 TLGDKLKNIGGLDKAGIGLDIVISGLSGATPAALVLADKNASTAKKYGAGFELANOVVNT 247
OY 351 TKAVSSYIIAQRVAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAEERF 410
DB 248 TKAVSSYIIAQRVAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAEERF 307
411 KKLGYDGNLLAEYORGTGTIDASVTAINTALAAIGVSAASAAA----- 454
DB 308 KKLGYDGNLLAEYORGTGTIDASVTAINTALAAIGVSAASAAAGSVIASPALLVSGIT 367
455 ----- 454
DB 368 GVISTILQYSKOAMFEHVANKIHNKIVEMEKNNHKNYFENGYDARYLANIQQNMKFLN 427
455 ----- 454
DB 428 LNKELQERVIATQQQWNNIGDLAGISRLGKRVLSKRAYVDAFEBGKHAKDLVOLD 487
455 ----- 454
DB 488 SANGIIVDSNSGAKTOHILFRPLPLPGTEHREVRQTKYEYITTKLINRVDSWKITDG 547
455 ----- 454

```

```

Db 548 AASTFDLNNVQRIEIDNAGNVTKTETKIAGLGGDNVFGSGTEIDGEGYD 607
QY 455 ----- 454
Db 608 RVHYSRGNYGALTTDATKETEGSYTVNREVTGKALHEVSTHVALGNREKEIEYRHS 667
QY 455 ----- 454
Db 668 NNQHHAGYTKDTLAVEEIIIGTSHNDIFKSGKFNDAFNGGCVDTIDGNGDNRLFSGK 727
QY 455 ----- 454
Db 728 GDDIIDGNGDDFIDGKGNDDLHGKGGDIFVHRKGGNDIITDSGNDKLSFSDSNLK 787
QY 455 ----- 454
Db 455 DLFEEKVHNLYITNSKKEKVTIOWMFEADPAKEVPYKATKDEKIEIITGONERITS 514
Db 788 DLFEEKVHNLYITNSKKEKVTIOWMFEADPAKEVPYKATKDEKIEIITGONERITS 847
QY 515 KOVDLIAGNGKITODELSKYVDVYELLKHSKNYNSLDKLSSVSFTSSNDSRNLY 574
Db 848 KOVDLIAGNGKITODELSKYVDVYELLKHSKNYNSLDKLSSVSFTSSNDSRNLY 907
QY 575 APTSMLDOSLSLOFARGS 593
Db 908 APTSMLDOSLSLOFARGS 926

```

# RESULT 15

US-07-779-171-8

Sequence 8, Application US/07779171

GENERAL INFORMATION:

APPLICANT: Potter, Andrew A.

APPLICANT: Redmond, Mark J.

APPLICANT: Hughes, Huv P.A.

TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster

STREET: 545 Middlefield Road, Suite 200

CITY: Menlo Park

STATE: California

COUNTRY: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/07/779,171

APPLICATION NUMBER: US/07/779,171

FILING DATE: 19911016

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Robins, Roberta L.

REGISTRATION NUMBER: 33,208

REFERENCE/DOCKET NUMBER: 29310-2001600

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-327-7250

TELEFAX: 415-327-2951

TELEX: 706141

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 951 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-779-171-8

Query Match 59.2%; Score 2126; DB 3; Length 951;  
Best Local Similarity 52.6%; Pred. No. 1.3e-162;

```

Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;
QY 111 SFPKTKAKKIIILYIPQNYQYDTEQNGLODLYKAAEELGIEVQRRERNRIATAQTSIGTI 170
Db 8 SFPKTKAKKIIILYIPQNYQYDTEQNGLODLYKAAEELGIEVQRRERNRIATAQTSIGTI 67
QY 171 QTAIGLTERGIYLSAPQIDKLLQKTRKAGALGASASIYQNAKAKTVISGIIISGYLA 230
Db 68 QTAIGLTERGIYLSAPQIDKLLQKTRKAGALGASASIYQNAKAKTVISGIIISGYLA 127
QY 231 GMDLDEALONNSNOHALAKAGLELTNSLENIANSVKTLDEGEQISQFSKLONIKIG 290
Db 128 GMDLDEALONNSNOHALAKAGLELTNSLENIANSVKTLDEGEQISQFSKLONIKIG 187
QY 291 TLGDKLKNIGLDKAGLGIDVYISGLSGATAALVADKNASTAKKVGAGFELANVGN 350
Db 188 TLGDKLKNIGLDKAGLGIDVYISGLSGATAALVADKNASTAKKVGAGFELANVGN 247
QY 351 TKAVSSYIIAORVAAGLSSTGPVALLIATSVSLATISPLAFAGIADKPFHAKSLESAER 410
Db 248 TKAVSSYIIAORVAAGLSSTGPVALLIATSVSLATISPLAFAGIADKPFHAKSLESAER 307
QY 411 KRLGYDGNLAEYQRTGTTIDASYTAINTALAAIAGVSAANA----- 454
Db 308 KRLGYDGNLAEYQRTGTTIDASYTAINTALAAIAGVSAANAAGSVIATALLVSGIT 367
QY 455 ----- 454
Db 368 GVISTIIQYSKQAMFEHVANKIHNKIVEMKNNHGNFYENGYDARYLANLODNMKFLLN 427
QY 455 ----- 454
Db 428 LNKELQAEKVIAITQOQNDNINIGDLAGISRLGEKVLGSKAYVDAFEEGKHIAKLVOLD 487
QY 455 ----- 454
Db 488 SANGIIVNSGKAKTQHILFRTPLTPGTEHREVRQYTGKEYITKLNINRYDSKITDG 547
QY 455 ----- 454
Db 548 AASTFDLNNVQRIEIDNAGNVTKTETKIAGLGGDNVFGSGTEIDGEGYD 607
QY 455 ----- 454
Db 608 RVHYSRGNYGALTTDATKETEGSYTVNREVTGKALHEVSTHVALGNREKEIEYRHS 667
QY 455 ----- 454
Db 668 NNQHHAGYTKDTLAVEEIIIGTSHNDIFKSGKFNDAFNGGCVDTIDGNGDNRLFSGK 727
QY 455 ----- 454
Db 728 GDDIIDGNGDDFIDGKGNDDLHGKGGDIFVHRKGGNDIITDSGNDKLSFSDSNLK 787
QY 455 DLFEEKVHNLYITNSKKEKVTIOWMFEADPAKEVPYKATKDEKIEIITGONERITS 514
Db 788 DLFEEKVHNLYITNSKKEKVTIOWMFEADPAKEVPYKATKDEKIEIITGONERITS 847
QY 515 KOVDLIAGNGKITODELSKYVDVYELLKHSKNYNSLDKLSSVSFTSSNDSRNLY 574
Db 848 KOVDLIAGNGKITODELSKYVDVYELLKHSKNYNSLDKLSSVSFTSSNDSRNLY 907
QY 575 APTSMLDOSLSLOFARGS 593
Db 908 APTSMLDOSLSLOFARGS 926

```

Search completed: March 2, 2001, 10:59:14  
Job time: 384 sec

**THIS PAGE BLANK (USPTO)**

04 of: US-09-306-689-12 to: A.Geneseq\_36.\* out\_format : pfs  
Date: Mar 2, 2001 10:31 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frim+n2p.model -DEV=xip  
-O=gcgn2.1/USPRO.spool/US09306689/rmat\_02032001\_102824\_9522/app-query.fasta.1.2389  
-DB=A.Geneseq\_36 -QFMT=fastan -SUFFIX=rag -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=dlsum62  
-TRANS=human4.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -MINLEN=0 -MAXLEN=2000000000  
-USER=US09306689\_@CGN1\_1187 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
-WAIT -THREADS=1

Search information block:

Query: US-09-306-689-12  
Query Length: 2088  
Database: A.Geneseq\_36.\*  
Database sequences: 268485  
Database length: 34193795  
Search time (sec): 109.740000

score\_list:

Sequence	Strd	Orig	ZScore	EScore	len	Documentation	8.7e-27	758	1	En
/SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT:R1187 + 380.50					576.12	5.7e-25	323	1	PI	

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT:R86998

seq\_documentation block:

ID R86998 standard; Protein: 758 AA.

AC R86998;

DT 04-JUL-1996 (first entry)

DE Enterohaemorrhagic E.coli hlyA gene product.

KW Enterohaemorrhagic Escherichia coli; virulent; EHEC; O157:H7 serotype;

KM detection; probe; primer; hlyA gene; enterohaemorrhagic colitis;

KW haemolytic uremic syndrome; mesenteric adenitis.

OS Escherichia coli (enterohaemorrhagic).

XX US5475098-A.

XX 12-DEC-1995.

XX 14-JUN-1994; 94US-0258188.

XX 14-JUN-1994; 94US-0258188.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Hall RH., Xu JG;

XX WPI: 1996-048546/05.

XX DR N-PSDB: T08098.

XX PT Enterohaemorrhagic E. coli (EHEC) nucleic acid sequences - useful  
XX for probe and primer design for sensitive and specific detection of  
XX EHEC  
XX Claim 1; Columns 37-42; 32pp; English.

XX Enterohaemorrhagic E.coli (EHEC) associated with enterohaemorrhagic  
CC colitis, haemolytic uremic syndrome and mesenteric adenitis have  
CC been found to carry a hlyA gene and a hlyB gene, separated by an  
CC intergenic region. The hly genes and the intergenic region are  
CC absent from bacteria not associated with these diseases and so  
CC provide a useful target for detecting EHEC pathogens, esp. O157:H7  
CC serotype E.coli. The present sequence is that of the protein  
CC encoded by the EHEC hlyA gene.

XX Sequence 758 AA;

alignment\_scores:

Quality:	403.00	Length:	395
Ratio:	1.807 <td>Gaps:</td> <td>11</td>	Gaps:	11
Percent Similarity:	56.456	Percent Identity:	30.633

alignment block:

US-09-306-689-12 x R86998

Align seg 1/1 to: R86998 from: 1 to: 758

964	GCACCTTGACTTGCAGATAAAATGCTTCACAGCTAAAGAGTGGTGC	1013
1	SerpheilleuenglyasnSeraspahisthnglyThrlyalaalal	17
1014	GGGTTTGAATTGGCAACCAAGTGTGTGTAATATTACCAAGCCGTTT	1063
17	aglyilegileuethrThrghlvalleuenglyasnvalglylyalvalis	34
1064	CTTCTTACATTTTATAGCCGAGCTTGCAGAGCTTATTCATCACTGGG	1113
34	ecelnglyrleuvalaglnarqmetalainglyleuSerThrtrala	50
1114	CCTGGGCTGCTTATTTATGCTTCTACTGTTCTGTCGATTAAGCCATT	1163
51	AlaserAlaglyleuethrSerAlaValMetleuAlaileSerProle	67
1164	AGCATTTGCCGGTATTCGCGATTAATTTATCATGCAAAAGTTAGAGA	1213
67	uSerpheleuAlaAlaAlaAspLysphegluargalalysglnleu	84
1214	GTTATGCCGAGCGCTTAAATAATAGCTATGACGAGATATATTATTA	1263
84	ertyrSerghluarphelyLysleuAsnTyrglnuLyspalaLeu	100
1264	GCAGATATATCAGCGGAGACGAGACTATGATGATGATGATGATGAT	1313
101	AlaAlaPheHisLysgluThrghlAlaileaspAlaAlaLeuThrTrl	117
1314	TAAATACGATTTGCCGCTATTTGCTGCTGCTGCTGCTGCTGACGCC	1362
117	eAsnThrValLeuSerSerValSerAlaValSerAlaAlaSerSera	134
1362	.....	1362
134	laserleuilegylalaproilleSerMetleuValSerAlaLeuThrly	150
1363	.....GATTTACATTTGAAAGT	1382
151	ThrIleSerGlyIleleuGlnuAlaSerLysGlnAlaMetPhegluHisVa	167
1383	TAAACATAATCTTGTATCAGCAATAGCAAAAGAGAGTGCACCATTC	1432
167	lAglulysPheAlaAlaargileasnclutrpGluLysglnHISGLYL	184
1433	AAACTGTTCCGAGAG	1449
184	YasnTyrPhegluasnnglyTyraSpAlaArgHisAlaAlaPheleuGlu	200
1450	.....GCTGATTTTGTCTAAAGAGTGCCTAATATTATA	1481



```
1663 .....AGTTAATCTCATCTAGTACATTACCTCG 1695
      ||| .....|||
119 LAlAlAlAlValGlyAlAlAlAlAlAlGlyLeIleThrGlyTyrThrSer 135
      : : : : :
1696 TCTAATGATTGAGAAAT .....GTAAT 1718
      : : : : :
136 AsnSerAlaGlyAlaThrLeuGlyValGlyLeuAlaAlaGlyLeuValGly 152
      : : : : :
1719 AGTGGCTCAACTTCATGTTGATCA .....AGTTATCTT 1756
      : : : : :
152 yMeAlaAlaAspAlaMetValGluAspValAsnTyrThrMetIleThrA 169
      : : : : :
1757 CTCTCAATTGCT ..... 1770
      : : : : :
169 spValGlnIleAlaGluArgThrLysAlaThrValThrThrAspAsnVal 185
      : : : : :
1771 .....AGGGATCTGACA 1784
      : : : : :
186 AlAlAlAlAlArgGlnGlyThrSerGlyAlaLysIleGlnThrSerGlnH 202
      : : : : :
1785 TTGGAGCTACGGCCTGCGCCTGGACGCGTCTCAAGATTGAGCTACG 1834
      ||| .....|||
202 sTrpSerTyrGlyLeuArgProGly .....GlnHsTrpSerTyrG 216
      : : : : :
1835 GCCTGGCTCCGGGTGGCTCTAGCCAGCATTTGAGCTACGCGCCTGGCCCT 1884
      ||| .....|||
216 LLeuArgProGly .....GlnHsTrpSerTyrGlyLeuArgPro 229
      : : : : :
1885 GGCAGCGGTAGCCAGATTGAGCTACGGCCTGCGCTCCGGGTGATCTCA 1934
      ||| .....|||
230 Gly .....GlnHsTrpSerTyrGlyLeuArgProGly .....Gln 241
      : : : : :
1935 GCATTGAGCTACGGCCTGCGCCTGGACGCGTCTCAAGATTGAGCT 1984
      : : : : :
241 uHsTrpSerTyrGlyLeuArgProGly .....GlnHsTrpSerT 255
      : : : : :
1985 ACGGCTGGCTCCGGGTGGCTCTAGCCAGCATTTGAGCTACGGCCTGGCG 2034
      ||| .....|||
255 yGlyLeuArgProGly .....GlnHsTrpSerTyrGlyLeuArg 268
      : : : : :
2035 CCTGGCAGCGGTAGCCAGATTGAGCTACGGCCTGCGCTCCGGGT 2079
      ||| .....|||
269 ProGly .....GlnHsTrpSerTyrGlyLeuArgProGly 280
```

**THIS PAGE BLANK (USPTO)**









DB 128 GMDLDEALONNSQNALAKAGELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 187  
QY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATPAALVLDKNASTAKKYGAGFELANOVYGN 350  
DB 188 TLGDKLKNIGGLDKAGLDIVISGLSGATPAALVLDKNASTAKKYGAGFELANOVYGN 247  
QY 351 TRAVSSYIIAORVAAGLSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 410  
DB 248 TRAVSSYIIAORVAAGLSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 307  
QY 411 KRIGYGDNLIAEYQGTGTIDASVTAINTALAAIAGVSAANA---DLTFEKVKHNLVI 467  
DB 308 KRIGYGDNLIAEYQGTGTIDASVTAINTALAAIAGVSAANA---DLTFEKVKHNLVI 367  
QY 468 TNSKKKVVIIQNMFRADPAKEVPNKATKDEKIEEIIIGONGERITSKOVDLLIANGNK 527  
DB 368 TNSKKKVVIIQNMFRADPAKEVPNKATKDEKIEEIIIGONGERITSKOVDLLIANGNK 427  
QY 528 ITODELSKVVDNVELLKHSKNTNSLDKLISVSATSSNDSRNVLVAPTSMLOSLSSL 587  
DB 428 ITODELSKVVDNVELLKHSKNTNSLDKLISVSATSSNDSRNVLVAPTSMLOSLSSL 487  
QY 588 QFARGSQHWSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSS 644  
DB 488 QFARGSQHWSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSS 544

## RESULT 5

US-08-878-748-10  
Sequence 10, Application US/08878748  
Patent No. 5969126  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/878,748  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016, 21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-878-748-10

Query Match 73.6%; Score 2642.5; DB 2; Length 544;  
Best Local Similarity 99.4%; Pred. No. 8,9e-192;  
Matches 534; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 111 SPFKTAKKIIILYIPONYQYDFQNGLODLVKAAELEIEVOREERNINATATQTSLSGT 170  
DB 8 SPFKTAKKIIILYIPONYQYDFQNGLODLVKAAELEIEVOREERNINATATQTSLSGT 67  
QY 171 QTAIGLTERGIYLSAPQIDKLQKTAKGALSAESIYONANKAKTIVLSIGISLSVLA 230  
DB 68 QTAIGLTERGIYLSAPQIDKLQKTAKGALSAESIYONANKAKTIVLSIGISLSVLA 127  
QY 231 GMDLDEALONNSQNALAKAGELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 290  
DB 128 GMDLDEALONNSQNALAKAGELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 187  
QY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATPAALVLDKNASTAKKYGAGFELANOVYGN 350  
DB 188 TLGDKLKNIGGLDKAGLDIVISGLSGATPAALVLDKNASTAKKYGAGFELANOVYGN 247  
QY 351 TRAVSSYIIAORVAAGLSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 410  
DB 248 TRAVSSYIIAORVAAGLSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 307  
QY 411 KRIGYGDNLIAEYQGTGTIDASVTAINTALAAIAGVSAANA---DLTFEKVKHNLVI 467  
DB 308 KRIGYGDNLIAEYQGTGTIDASVTAINTALAAIAGVSAANA---DLTFEKVKHNLVI 367  
QY 468 TNSKKKVVIIQNMFRADPAKEVPNKATKDEKIEEIIIGONGERITSKOVDLLIANGNK 527  
DB 368 TNSKKKVVIIQNMFRADPAKEVPNKATKDEKIEEIIIGONGERITSKOVDLLIANGNK 427  
QY 528 ITODELSKVVDNVELLKHSKNTNSLDKLISVSATSSNDSRNVLVAPTSMLOSLSSL 587  
DB 428 ITODELSKVVDNVELLKHSKNTNSLDKLISVSATSSNDSRNVLVAPTSMLOSLSSL 487  
QY 588 QFARGSQHWSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSS 644  
DB 488 QFARGSQHWSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSS 544

## RESULT 6

US-09-124-491-10  
Sequence 10, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996

```

APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-10

```

```

Query Match          73.6%; Score 2642.5; DB 3; Length 544;
Best Local Similarity 99.4%; Pred. No. 8.9e-192;
Matches 534; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```

```

OY 111 SEPKTGAKKILLYIPONTQYDTEQNGIQLDYLKAAEELGIEVQREERNINATAOTSIGTI 170
    |||||||
DB 8 SEPKTGAKKILLYIPONTQYDTEQNGIQLDYLKAAEELGIEVQREERNINATAOTSIGTI 67
OY 171 QTAIGLTERGIVLSAPQIDKLQKTKAGQALGSAESIYONANKAKTVLSGISIIGSVLA 230
    |||||||
DB 68 QTAIGLTERGIVLSAPQIDKLQKTKAGQALGSAESIYONANKAKTVLSGISIIGSVLA 127
OY 231 GMDLDEALONNSNOHALKAGLELTNSLIENIANSVKTLDFEGEQISQFGSKLQNIKIGL 290
    |||||||
DB 128 GMDLDEALONNSNOHALKAGLELTNSLIENIANSVKTLDFEGEQISQFGSKLQNIKIGL 187
OY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATAALVLAADKNASTAKKVGAFELANQVGN 350
    |||||||
DB 188 TLGDKLKNIGGLDKAGLDIVISGLSGATAALVLAADKNASTAKKVGAFELANQVGN 247
OY 351 TKAVSSYLLAORVAAGLSSTGPVALLASTVSLAISPLAFAGIADKFHNAKLSLEYAERF 410
    |||||||
DB 248 TKAVSSYLLAORVAAGLSSTGPVALLASTVSLAISPLAFAGIADKFHNAKLSLEYAERF 307
OY 411 KKLGYDGNLLAEYORGTGTIDASTVTAINTALAAIAGGVSAAAA---DLTFEKVKNHLYI 467
    |||||||
DB 308 KKLGYDGNLLAEYORGTGTIDASTVTAINTALAAIAGGVSAAAA---DLTFEKVKNHLYI 367
OY 468 TNSKREKVTIOWMFREADFAKEVPYKATKDEKIEIIGONGERTTSKQVODLLIAKNGK 527
    |||||||
DB 368 TNSKREKVTIOWMFREADFAKEVPYKATKDEKIEIIGONGERTTSKQVODLLIAKNGK 427
OY 528 ITQDLSLVVNYNELKLSKNTNSLDKLISVSFAFTSSNDRNVLPVPTSLDQSLSSL 587
    |||||||
DB 428 ITQDLSLVVNYNELKLSKNTNSLDKLISVSFAFTSSNDRNVLPVPTSLDQSLSSL 487
OY 588 OFARSSQHSYGLRPGSSQDMSTYGLRPGSSQHSYGLRPGSSQDMSTYGLRPGSS 644
    |||||||
DB 488 OFARSSQHSYGLRPGSSQDMSTYGLRPGSSQHSYGLRPGSSQDMSTYGLRPGSS 544

```

```

RESULT 7
US-08-387-156-8
Sequence 8, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUW P. A.

```

```

TITLE OF INVENTION: GARRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-8

```

```

Query Match          67.6%; Score 2426; DB 1; Length 977;
Best Local Similarity 55.1%; Pred. No. 4.9e-175;
Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

```

OY 111 SEPKTGAKKILLYIPONTQYDTEQNGIQLDYLKAAEELGIEVQREERNINATAOTSIGTI 170
    |||||||
DB 8 SEPKTGAKKILLYIPONTQYDTEQNGIQLDYLKAAEELGIEVQREERNINATAOTSIGTI 67
OY 171 QTAIGLTERGIVLSAPQIDKLQKTKAGQALGSAESIYONANKAKTVLSGISIIGSVLA 230
    |||||||
DB 68 QTAIGLTERGIVLSAPQIDKLQKTKAGQALGSAESIYONANKAKTVLSGISIIGSVLA 127
OY 231 GMDLDEALONNSNOHALKAGLELTNSLIENIANSVKTLDFEGEQISQFGSKLQNIKIGL 290
    |||||||
DB 128 GMDLDEALONNSNOHALKAGLELTNSLIENIANSVKTLDFEGEQISQFGSKLQNIKIGL 187
OY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATAALVLAADKNASTAKKVGAFELANQVGN 350
    |||||||
DB 188 TLGDKLKNIGGLDKAGLDIVISGLSGATAALVLAADKNASTAKKVGAFELANQVGN 247
OY 351 TKAVSSYLLAORVAAGLSSTGPVALLASTVSLAISPLAFAGIADKFHNAKLSLEYAERF 410
    |||||||
DB 248 TKAVSSYLLAORVAAGLSSTGPVALLASTVSLAISPLAFAGIADKFHNAKLSLEYAERF 307
OY 411 KKLGYDGNLLAEYORGTGTIDASTVTAINTALAAIAGGVSAAAA-----+ 454
    |||||||
DB 308 KKLGYDGNLLAEYORGTGTIDASTVTAINTALAAIAGGVSAAAA-----+ 367
OY 455 ----- 454
DB 368 GVSTIIQYSKQMEHVAHNRKIHNVKEMERNNHKNYFENGVDARYLANQDMKFLIN 427
OY 455 ----- 454

```



|||||  
Db 908 APTSMLODSSLQFARSGQHSYGLRPGSGSQDMSYGLRPGSGSQDMSYGLRPGSGSQD 967  
QY 635 WSYGLRPGGS 644  
|||||  
Db 968 WSYGLRPGGS 977

## RESULT 9

US-08-878-748-8  
Sequence 8, Application US/08878748  
Patent No. 5969126  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/878,748  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-878-748-8

Query Match 67.68; Score 2426; DB 2; Length 977;  
Best Local Similarity 55.18; Pred. No. 4.9e-175;  
Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

QY 111 SFPKGAKKIILYIPONTQYOTEGONGIÖDLYKAAEEIGIEYÖRERNINATAGTSLGTI 170  
|||||  
Db 8 SFPKGAKKIILYIPONTQYOTEGONGIÖDLYKAAEEIGIEYÖRERNINATAGTSLGTI 67  
QY 171 QTAIGLTERGIYLSAPÖIDKLÖKTRKAGÖLSAESIYÖNANKAKTVLSGIOSIIGSVILA 230  
|||||  
Db 68 QTAIGLTERGIYLSAPÖIDKLÖKTRKAGÖLSAESIYÖNANKAKTVLSGIOSIIGSVILA 127  
QY 231 GMDLDEALÖNNSNOHALAKAGIELTNSLIENIANSVKTLDEGEQISÖFGSKLÖNIKIGL 290  
|||||  
Db 128 GMDLDEALÖNNSNOHALAKAGIELTNSLIENIANSVKTLDEGEQISÖFGSKLÖNIKIGL 187

QY 291 TLGDKLKNIGGLDKAGLGLDVIISGLSGATAALVLDKNASTAKVKGAFELANOVGNI 350  
|||||  
Db 188 TLGDKLKNIGGLDKAGLGLDVIISGLSGATAALVLDKNASTAKVKGAFELANOVGNI 247  
QY 351 TRAVSSYILAO RVAAGLSSTGPVALLIASTVSLAISPLAFAGIADKFHNASLSESAERF 410  
|||||  
Db 248 TRAVSSYILAO RVAAGLSSTGPVALLIASTVSLAISPLAFAGIADKFHNASLSESAERF 307  
QY 411 KKLGYDGNLLAEYORGTGTIDASVTAINTAALAIAGVSAAAA----- 454  
|||||  
Db 308 KKLGYDGNLLAEYORGTGTIDASVTAINTAALAIAGVSAAAAASVIAISPILLVSGIT 367  
QY 455 ----- 454  
Db 368 GVISTITLOYSKÖAMFEHVANKIHNKIVEMKNNHKNYFENGYDARYLANIÖDNMKFLN 427  
QY 455 ----- 454  
Db 428 LNKELÖAERVAITÖQÖMÖNNIGDLAGISRLGEKVLISKAVYDAFEBSKHIAKDLVÖLD 487  
QY 455 ----- 454  
Db 488 SANGIIVDSNGKAKTÖHILFRTPLTPGTEHREYÖTGKEYITKLNINRVDSWKITDG 547  
QY 455 ----- 454  
Db 548 AASTFDLTNNVÖRIGIELDNAGNVTKETKRIIAKLGEÖDNVFGSGTTEIDGEGYD 607  
QY 455 ----- 454  
Db 608 RVHYSRGVAGALTIDATETEGSGTYVVRPEYETGALHEVSTHTALVGNREKIEYRHS 667  
QY 455 ----- 454  
Db 668 NNOHHAGYTTÖDLKAVEIIGTSHNDIFKSGKFNDAPFNGDGYDTIDGNDGNRLFGGK 727  
QY 455 ----- 454  
Db 728 GDDIILGNGDDFDGKGNDLHGKGDDIFVHRKGDNDIITDSÖDNKLSFSÖSNLK 787  
QY 455 ----- 454  
Db 455 ----- 454  
QY 455 ----- 454  
Db 788 DTFEKKVKNHIVTNSKKEKVTIQNMFREADFAKEVPYKATKDEKIEIIGÖNERITS 514  
|||||  
Db 788 DTFEKKVKNHIVTNSKKEKVTIQNMFREADFAKEVPYKATKDEKIEIIGÖNERITS 847  
QY 515 KÖVDDLIAKNGKITÖDELISKVVDNYELLKHSKNVTNSLDKLISSVSAFTSSNDSRNLY 574  
|||||  
Db 848 KÖVDDLIAKNGKITÖDELISKVVDNYELLKHSKNVTNSLDKLISSVSAFTSSNDSRNLY 907  
QY 575 APTSMLODSSLQFARSGQHSYGLRPGSGSQDMSYGLRPGSGSQDMSYGLRPGSGSQD 634  
|||||  
Db 908 APTSMLODSSLQFARSGQHSYGLRPGSGSQDMSYGLRPGSGSQDMSYGLRPGSGSQD 967  
QY 635 WSYGLRPGGS 644  
|||||  
Db 968 WSYGLRPGGS 977

RESULT 10  
US-09-124-491-8  
Sequence 8, Application US/09124491  
Patent No. 6022960

GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA

ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-8

Query Match 67.6%; Score 2426; DB 3; Length 977;  
Best local Similarity 55.1%; Pred. No. 4.9e-175;  
Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

QY 111 SPKGTAKKIIILYIPONYQYDTEQNGLODLVKAEBELGIEVOREERNINATATQTSIGTI 170  
DB 8 SPKGTAKKIIILYIPONYQYDTEQNGLODLVKAEBELGIEVOREERNINATATQTSIGTI 67  
QY 171 QTAIGLTERGIVASAPQIDLLQKTRAGALGSAESIYONANKAKTVLSGIQSLGSVLA 230  
DB 68 QTAIGLTERGIVASAPQIDLLQKTRAGALGSAESIYONANKAKTVLSGIQSLGSVLA 127  
QY 231 GMDLDELQNSNOHAKAGLELTNSLIENIANSVKTTDEFEQJISQFSGSKLQNIKGLG 290  
DB 128 GMDLDELQNSNOHAKAGLELTNSLIENIANSVKTTDEFEQJISQFSGSKLQNIKGLG 187  
QY 291 TLCDKLKNIIGDLKAGIGLDIVISGLSGATAVLADKNASTAKKYGAGELANOVYNT 350  
DB 188 TLCDKLKNIIGDLKAGIGLDIVISGLSGATAVLADKNASTAKKYGAGELANOVYNT 247  
QY 351 TKAVSSYTLAQRVAAGISSTGPAALIASVSLAISPLAFAGIADKNHAKSLSEVAERF 410  
DB 248 TKAVSSYTLAQRVAAGISSTGPAALIASVSLAISPLAFAGIADKNHAKSLSEVAERF 307  
QY 411 KKLGDNDNLAEYORGTGTIDASVTAINATALAAGVSAANA----- 454  
DB 308 KKLGDNDNLAEYORGTGTIDASVTAINATALAAGVSAANAAGVSIAPILALVSGIT 367  
QY 455 ----- 454  
DB 368 GVIATLLQYSQAMFEHVANKIHKKIVEMKNNHKKYFENGYDARYLANLQDNMFLN 427  
QY 455 ----- 454  
DB 428 LNKELQAEVIAITQOOWNNIGDLAGISRLGKVLGSKAYVDAFEEGHKIKADKLVLQD 487

QY 455 ----- 454  
DB 488 SANGIIDVNSGKAKTQHILFRTPLTPGHEHREVOTGXEYITKLNINRVDSMKITDG 547  
QY 455 ----- 454  
DB 548 AASTFDLNNVORIGIELDNAGNVTKRETKLIAKLGRDDNVFVSGTTEIDGGEYD 607  
QY 455 ----- 454  
DB 608 RVHSGNNGALTIDATKETEGOSTYVNRFEVTKALHEVTSHTALVGNREKIEYRIS 667  
QY 455 ----- 454  
DB 668 NQGHAGYTKDTLKAVEELIGTSHNDIFKSGFNDAFNGDGVDTIDGNDRLFGK 727  
QY 455 ----- 454  
DB 728 GDDILDGNGDDFIDGKGNDDLHGKRGDDIFVHRKGDNDITDSGNDKLSFSDSNLK 787  
QY 455 DLFEKVKHNLVTNKKKEKVTIONMFEADFAKEYPNKATKDEKIEELIGONGERITS 514  
DB 788 DLFEKVKHNLVTNKKKEKVTIONMFEADFAKEYPNKATKDEKIEELIGONGERITS 847  
QY 515 KQVDDLKANGKKTODELSKVVDNTELLKHSKNVTNSLDKLISVSAAFTSSNDSRNVLY 574  
DB 848 KQVDDLKANGKKTODELSKVVDNTELLKHSKNVTNSLDKLISVSAAFTSSNDSRNVLY 907  
QY 575 APTSMDDQSLSLQFARGSHWSYGLRPGSGSDMSYGLRPGSSQSHWSYGLRPGSGSD 634  
DB 908 APTSMDDQSLSLQFARGSHWSYGLRPGSGSDMSYGLRPGSSQSHWSYGLRPGSGSD 967  
QY 635 WSYGLRPGGS 644  
DB 968 WSYGLRPGGS 977

RESULT 11  
US-08-455-970A-12  
Sequence 12, Application US/08455970A  
Patent No. 5708155  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P.A.  
TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN  
TITLE OF INVENTION: CHIMERAS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,970A  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 327-3400



TELEFAX: (415) 327-3231  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 936 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-455-970A-12

Query Match 61.0%, Score 2189, DB 1; Length 936;  
 Best Local Similarity 53.1%; Pred. No. 3,7e-157;  
 Matches 493; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

QY 111 SPPKTKAKIILYIPONYOYDEONGLODLYKAEELGIEVQREERNINIAAOTSIGTI 170
DB 8 SPPKTKAKIILYIPONYOYDEONGLODLYKAEELGIEVQREERNINIAAOTSIGTI 67
QY 171 QTAIGLTERGIVLSAPQIDKLLQTKAGALGSASIVQNNANKATVLSIGIOSIIGSVLA 230
DB 68 QTAIGLTERGIVLSAPQIDKLLQTKAGALGSASIVQNNANKATVLSIGIOSIIGSVLA 127
QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 290
DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 187
QY 291 TLGDKLKNIGIGLDRKGLGLDVIISGLSGATAALVADKNASTAKKVGAGFELANOVGNI 350
DB 188 TLGDKLKNIGIGLDRKGLGLDVIISGLSGATAALVADKNASTAKKVGAGFELANOVGNI 247
QY 351 TKAVSSYLLAORVAAGLSSTGPVALLIASTVSLAISPPLAFAGIADKFHNAKSLSYSYAEERF 410
DB 248 TKAVSSYLLAORVAAGLSSTGPVALLIASTVSLAISPPLAFAGIADKFHNAKSLSYSYAEERF 307
QY 411 KKLGYDGNLLAEYQRTGTIDASYTAINTALAAIAGVSAANA----- 454
DB 308 KKLGYDGNLLAEYQRTGTIDASYTAINTALAAIAGVSAANAASPILALVSGIT 367
QY 455 ----- 454
DB 368 GVISTIILOYSKOAMFEHVANKIHKIVEMKNNHKNYFENGCDARYLANQDNKKFLIN 427
QY 455 ----- 454
DB 428 LNKELQAEIRVIAITQQQKNDNINIGDLAGISRLGEKYLSGKAYVDAPEEGKHIAKDLVQLD 487
QY 455 ----- 454
DB 488 SANGIIDVNSGKAKTOHILFRTPLLPTEHREHREVOTGKVEYITKLNINRYDSMKITDG 547
QY 455 ----- 454
DB 548 AASSTFDLTNNVQRIIGIELDNAGNVTKTKETKIIAKLGEGBDNVEVSGSTTEIDGEGYD 607
QY 455 ----- 454
DB 608 RVHYSRGNVGAITIDATKETEOGISTYVNRPEVTGKALHEVYSTHTALVGNREKIEYRHS 667
QY 455 ----- 454
DB 668 NNOHHAGYYTKDTLKAVEEIIIGTSHNDIFKSKFNDAFNGGDGVDITIDGNDGNDRLFGSK 727
QY 455 ----- 454
DB 728 GDDIIDGNGDDEFIDGKGNDLHGKGGDDIFVHRKGGDNDITDSDNDKLSFSDSNLK 787
QY 455 DLTEFKVHNLYVTNSKREKVTIQNMFPREADPAKEVPYKATKDKIEIIGONGERTIS 514
DB 788 DLTEFKVHNLYVTNSKREKVTIQNMFPREADPAKEVPYKATKDKIEIIGONGERTIS 847
QY 515 KOVDLLAKNGCKITQDELISKVVDNYELLKHSKNVTNSLDKLISVSASFSSNDGRNVLY 574
DB 848 KOVDLLAKNGCKITQDELISKVVDNYELLKHSKNVTNSLDKLISVSASFSSNDGRNVLY 907
  
```

QY 575 APTSMLDOSLSSLOFARGSOHWSYGLRPG 603  
 DB 908 APTSMLDOSLSSLOFARGSOHWSYGLRPG 936

RESULT 12  
 US-07-777-715-9

; Sequence 9, Application US/07777715

; Patent No. 5273889

; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew

; APPLICANT: Campos, Manuel

; APPLICANT: Hughes, Huw P. A.

; TITLE OF INVENTION: CYTOKINE-LEUKOTOXIN GENE FUSIONS AND

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESS: Morrison & Foerster

; STREET: 545 Middlefield Road, Suite 200

; CITY: Menlo Park

; STATE: California

; COUNTRY: USA

; ZIP: 94025

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/777,715

; FILING DATE: 19911016

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Robins, Roberta L.

; REGISTRATION NUMBER: 33,208

; REFERENCE/DOCKET NUMBER: 29310-2001320

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-327-7250

; TELEFAX: 415-327-2951

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1069 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-07-777-715-9

Query Match 59.3%, Score 2131, DB 1; Length 1069;

Best Local Similarity 52.6%; Pred. No. 1.1e-152;

Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

QY 111 SPPKTKAKIILYIPONYOYDEONGLODLYKAEELGIEVQREERNINIAAOTSIGTI 170
DB 8 SPPKTKAKIILYIPONYOYDEONGLODLYKAEELGIEVQREERNINIAAOTSIGTI 67
QY 171 QTAIGLTERGIVLSAPQIDKLLQTKAGALGSASIVQNNANKATVLSIGIOSIIGSVLA 230
DB 68 QTAIGLTERGIVLSAPQIDKLLQTKAGALGSASIVQNNANKATVLSIGIOSIIGSVLA 127
QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 290
DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 187
QY 291 TLGDKLKNIGIGLDRKGLGLDVIISGLSGATAALVADKNASTAKKVGAGFELANOVGNI 350
DB 188 TLGDKLKNIGIGLDRKGLGLDVIISGLSGATAALVADKNASTAKKVGAGFELANOVGNI 247
QY 351 TKAVSSYLLAORVAAGLSSTGPVALLIASTVSLAISPPLAFAGIADKFHNAKSLSYSYAEERF 410
DB 248 TKAVSSYLLAORVAAGLSSTGPVALLIASTVSLAISPPLAFAGIADKFHNAKSLSYSYAEERF 307
  
```

OY	411	KKLGVGDGNLLAEYORGTGTTIDASVTAINNTALAAIAGVSAAAA-----	454
Dd	308	KKLGVDGNLLEAFQRGITGTTIDASVTAINNTALAAIAGVSAAAAAGSVIASPIALLYSGIT	367
OY	455	----	454
Dd	368	GVISTITLOYSKOAMFEHVANKIHNKIVEMEKNNHGKNYPENGYPDARYLANLDONMKFLLN	427
OY	455	----	454
Dd	428	LNLQLAERVAIAITQQQWDDNNGDLAGISRLGEKVLSCGAAYDAFEGEGRHFKADKLVOYD	487
OY	455	----	454
Dd	488	SANGIIDVSNCGAKTOHIILFRPLLPCTHERBERVOTGKEYITTIKLINRVDSWKITDG	547
OY	455	----	454

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/777, 715  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/571, 301  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ROBINS, ROBERTA L.  
 REGISTRATION NUMBER: 33, 208  
 REFERENCE/DOCKET NUMBER: 9000-0013.21  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 617-8999  
 TELEFAX: (415) 327-3231  
 INFORMATION FOR SEQ. ID NO.: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1069 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-170-126-4

Query Match 59.3%; Score 2131; DB 1; Length 1069;  
 Best Local Similarity 52.6%; Pred. No. 1,1e-152;  
 Matches 486; Conservative 0; Mismatches 0; Indels 436; Gaps 1

OY	455	-----	454
Db	668	NQHHAGYTTKDTLKAVEEIIIGTSHNDIFKSGKFNDAFNGDGVDTIDGNDGDRLEFGK	727
OY	455	-----	454
Db	728	GDDILLDGGNGDFTIDGKGNDLHGKGDDI FVHRKGDNDIITDSDDNKLSPSDSNLK	787
OY	455	DTEFEKVHNLVITNSKKEKVTIONMWEADFAKEVBNYKATFDEKIEELIIGONGERITS	514
Db	788	DTEFEKVHNLVITNSKKEKVTIONMWEADFAKEVBNYKATFDEKIEELIIGONGERITS	847
OY	515	KQVDDLLAKNGKTTODELSKYVDNYELLKHKSNVNTSLDKLISSVAFSSSDSRNVLV	574
Db	848	KQVDDLLAKNGKTTODELSKYVDNYELLKHKSNVNTSLDKLISSVAFSSSDSRNVLV	907
OY	575	APTSMLDQSLSSLOPFARSGQ	594
Db	908	APTSMLDQSLSSLOPFARSGQ	927

RESULT 13  
 US-08-170-126-4  
 ; Sequence 4, Application US/08170126  
 ; Patent No. 5594107  
 ; GENERAL INFORMATION:  
 ; APPLICANT: POTTER, ANDREW  
 ; APPLICANT: CAMPOS, MANUEL  
 ; APPLICANT: HUGHES, HUM P. A.  
 ; TITLE OF INVENTION: CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES  
 ; THEREOF  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: REED & ROBINS  
 ; STREET: 635 BRYANT STREET  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: UNITED STATES OF AMERICA  
 ; ZIP: 94301  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/170,126  
 ; FILING DATE: 20-DEC-1993  
 ; CLASSIFICATION: 530

```

: PRIOR APPLICATION DATA: US 07/777,715
: APPLICATION NUMBER: US 07/777,715
: PRIOR APPLICATION DATA: US 07/571,301
: APPLICATION NUMBER: US 07/571,301
: ATTORNEY/AGENT INFORMATION:
: NAME: ROBINS, ROBERTA L.
: REGISTRATION NUMBER: 33,208
: REFERENCE/DOCKET NUMBER: 9000-0013.21
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 617-8999
: TELEFAX: (415) 327-3231
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1069 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-170-126-4

Query Match      59.3%; Score 2131; DB 1; Length 1069;
Best Local Similarity 52.6%; Pred. No. 1,1e-152;
Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps

OY 111 SPPKGAKKIILTYIPONTQVYTEGNGLODLYKAAEEIGIEVOEERENNATATQTSIGTI 1707
Db 8 SPPTGAKKIITLTPONTQVYTEGNGLODLYKAAEEIGIEVOEERENNATATQTSIGTI 67
OY 171 QTAIGLERGIVLSAPOLDIKLQRTKAGALGSNESIVONANKAKTVLSGIQSILGSVLA 2308
Db 68 QTAIGLERGVISLPAPQIDKLQRTKAGALGSNESIVONANKAKTVLSGIQSILGSVLA 1277
OY 231 GMDLDDEALONNSNOHALAKAGLELTNSLIENIANSVYTTLDFEGQISOFSCKLONIKGLG 2300
Db 231 GMDLDDEALONNSNOHALAKAGLELTNSLIENIANSVYTTLDFEGQISOFSCKLONIKGLG 1877
OY 291 TLGDPLKNIGGLDWAGLGADVISLLSGATAVLADKNASTAKKVAGFELANOVYGN 3500
Db 188 TLGDPLKNIGGLDWAGLGADVISLLSGATAVLADKNASTAKKVAGFELANOVYGN 2477
OY 351 TKAVSSYTLAORVAAGLSSTGPVALIAITSVSLAISPLAFAGIADKRNNHAKSLESYAERE 4100
Db 248 TKAVSSYTLAORVAAGLSSTGPVALIAITSVSLAISPLAFAGIADKRNNHAKSLESYAERE 3077
OY 411 KKLGYDGNLLAEYQRTGTIDASVTAINFLAAIAGGVSAHA----- 4544
Db 308 KKLGYDGNLLAEYQRTGTIDASVTAINFLAAIAGGVSAHAASYIASPIALLVSGIT 3677
OY 455 ----- 4544
Db 368 GVISTIIQYSKOAMEEHVANKIHKKIVEWERNHHGKNYPENGYDARYLANLQDNMKFLN 4227
OY 455 ----- 4544
Db 455 ----- 4544
OY 455 ----- 4544
Db 488 SANGIIDVSNKGAKAQHLIFRPLTPPGTEHREREVOTGKYEYITKLINRVDSWKITTDG 5474
OY 455 ----- 4544
Db 548 AASSTFDLTJNVVGRIEILDAGNAVTKETKIIAKLGEDDNVFGSGTTEIDGGEGYD 6077
OY 455 ----- 4544
Db 608 RVHTSRGNYGALLTDATKETEQSYTNRRVEYTGKAHEVYSTHTALVGNREKEIERYHS 6677
OY 455 ----- 4544
Db 668 NNQHAGYYTTDLTKAVBEIIGTSHNDIFKSGKENDAFNGGDGVDTIDGNDGNDRLFGGK 7277
OY 455 ----- 4544

```

Db 728 GDDILDCGNDPFDIGCKGNDLHGCGDDIFVHRKGGNDIITDSDGNDKLSFSDSNLK 787  
QY 455 DLTEFKVHNLYITNSKKKEKVTIONMFPREADFAKEVPYKATKDEKIEIIGONGERITS 514  
Db 788 DLTEFKVHNLYITNSKKKEKVTIONMFPREADFAKEVPYKATKDEKIEIIGONGERITS 847  
QY 515 KOVDDLAKNGKKTODELSKVVDNYELKHSKNVTNSLDKLSSVSFAFTSSNDSRNLY 574  
Db 848 KOVDDLAKNGKKTODELSKVVDNYELKHSKNVTNSLDKLSSVSFAFTSSNDSRNLY 907  
QY 575 APTSMLDQSLSSLOFARGSQ 594  
Db 908 APTSMLDQSLSSLOFARGSQ 927  
RESULT 14  
US-08-954-418-4  
; Sequence 4, Application US/08954418  
; Patent No. 6096320  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW  
APPLICANT: CAMPOS, MANUEL  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: CYTOKINE-CYTOFOXIN GENE FUSIONS AND USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/954,418  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/170,126  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/571,301  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9000-0013.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1069 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-954-418-4  
Query Match 59.38; Score 2131; DB 3; Length 1069;  
Best Local Similarity 52.68; Pred. No. 1,1e-152;  
Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps 1;  
QY 111 SEPTGAKKILYIPONQYDTGONGLODYVKAEEIGIEVOREERNINATAQTSIGTI 170  
Db 8 SEPTGAKKILYIPONQYDTGONGLODYVKAEEIGIEVOREERNINATAQTSIGTI 67  
QY 171 QTAIGTERGIVLSAPQIDKLLQKTKAGQALGSAESIYQNNANKKATVLSIGISLVIA 230

Db 68 QTAIGTERGIVLSAPQIDKLLQKTKAGQALGSAESIYQNNANKKATVLSIGISLVIA 127  
QY 231 GMDIDELALQNNNSQHALAKAGLELTNSLIENIANSVKTLDFEQDQISPGSKLONIKGLG 290  
Db 128 GMDIDELALQNNNSQHALAKAGLELTNSLIENIANSVKTLDFEQDQISPGSKLONIKGLG 187  
QY 291 TLGDKLNKINIGGLDRAKGLDVIYSGLSGATPAALVADNASTAKKVGAGFELANQVGN 350  
Db 188 TLGDKLNKINIGGLDRAKGLDVIYSGLSGATPAALVADNASTAKKVGAGFELANQVGN 247  
QY 351 TKAVSSYTLAORVAAGLSSTGPVAAALASTVSLAISPLAFAGIADRFNHAKSLESTARE 410  
Db 248 TKAVSSYTLAORVAAGLSSTGPVAAALASTVSLAISPLAFAGIADRFNHAKSLESTARE 307  
QY 411 KKLGYDGNLLAEYORGTTIDASTYNTINTLALAIAGVSAANA----- 454  
Db 308 KKLGYDGNLLAEYORGTTIDASTYNTINTLALAIAGVSAANAAGSVIASPLALLVSGIT 367  
QY 455 ----- 454  
Db 368 GYISTIILOYSKQAMFEHVANKIHNKIVWEKNHGNKFENGIDARYLANLQDNKKFLN 427  
QY 455 ----- 454  
Db 428 LNKELQARVIAITQQQMDNNIGLAGISRLGEKVLSCAKAYVDAFEEGKHAKKLVOLD 487  
QY 455 ----- 454  
Db 488 SANGIIDVNSNGAKAKOHLFPTPLPGTEHREVRVOTGKYEYITKLNINRDSMKITDG 547  
QY 455 ----- 454  
Db 548 AASTFDLTNNVQRIELDNAGNVTKETKTIKAGLEGDDNVFVSGTTEIDGEGSYD 607  
QY 455 ----- 454  
Db 608 RVHYSRGNYGALTIDATKETEQGSYTVNRFVETGALHEVSTHTALVGNREKIEYRHS 667  
QY 455 ----- 454  
Db 668 NNQHHGYTDTLKAVEIIGTSHNDIFKSGKFNDAFNGDGVDTIDGNDGNDRLFGK 727  
QY 455 ----- 454  
Db 728 GDDILDCGNDPFDIGCKGNDLHGCGDDIFVHRKGGNDIITDSDGNDKLSFSDSNLK 787  
QY 455 DLTEFKVHNLYITNSKKKEKVTIONMFPREADFAKEVPYKATKDEKIEIIGONGERITS 514  
Db 788 DLTEFKVHNLYITNSKKKEKVTIONMFPREADFAKEVPYKATKDEKIEIIGONGERITS 847  
QY 515 KOVDDLAKNGKKTODELSKVVDNYELKHSKNVTNSLDKLSSVSFAFTSSNDSRNLY 574  
Db 848 KOVDDLAKNGKKTODELSKVVDNYELKHSKNVTNSLDKLSSVSFAFTSSNDSRNLY 907  
QY 575 APTSMLDQSLSSLOFARGSQ 594  
Db 908 APTSMLDQSLSSLOFARGSQ 927  
RESULT 15  
US-07-908-253-2  
; Sequence 2, Application US/07908253  
; Patent No. 5534256  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: HARLAND, RICHARD J.  
TITLE OF INVENTION: HAEMOPHILUS SOMNUS OUTER MEMBRANE  
TITLE OF INVENTION: PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ROBERTA L. ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO

```

STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/908,253
FILING DATE: 19920702
CLASSIFICATION: 420
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9000-0026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 926 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-908-253-2

```

```

Query Match      59.2%; Score 2126; DB 1; Length 926;
Best Local Similarity 52.6%; Pred. No. 2.1e-152;
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

```

OY 111 SPFKTGAKKIIYIPONYOYDEQNGIIDLKYAAEELGIEYREBERNNIATQTSIGTT 170
    |||||||
DB 8 SPFKTGAKKIIYIPONYOYDEQNGIIDLKYAAEELGIEYREBERNNIATQTSIGTT 67
    |||||||
OY 171 QTAIGLTERGIYLSAPOIDKLQKTRAGALGSAESIYONANKAKTVLSGIGSVLA 230
    |||||||
DB 68 QTAIGLTERGIYLSAPOIDKLQKTRAGALGSAESIYONANKAKTVLSGIGSVLA 127
    |||||||
OY 231 GMDLDELQNSNQHALAAGLELTSLENTANSYKTIIDEPGEQISQPSKIQNTKGLG 290
    |||||||
DB 128 GMDLDELQNSNQHALAAGLELTSLENTANSYKTIIDEPGEQISQPSKIQNTKGLG 187
    |||||||
OY 291 TLGDKLKNIGGLDAGLGLDVISGLSGATAALVLDKNASTAKKVGAGFELANQVGN 350
    |||||||
DB 188 TLGDKLKNIGGLDAGLGLDVISGLSGATAALVLDKNASTAKKVGAGFELANQVGN 247
    |||||||
OY 351 TKAVSSYIIAQRYAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAE 410
    |||||||
DB 248 TKAVSSYIIAQRYAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAE 307
    |||||||
OY 411 KRLGYGDNLLEYORGTGTIDASYTAINTALAAAGVSAANA----- 454
    |||||||
DB 308 KRLGYGDNLLEYORGTGTIDASYTAINTALAAAGVSAANA----- 367
    |||||||
OY 455 ----- 454
DB 368 GVISTILQYSKQAMFEHVANKINKIVEMEKNNHGKNYPENGDAYRIANLQDNMKFLN 427
    |||||||
OY 455 ----- 454
DB 428 LNKELQAEVIAITQOQMDNNIGDLAGISRLGEKVLGKAYVDAFEGKHAKDKLVOLD 487
    |||||||
OY 455 ----- 454
DB 488 SANGIIDVNSGAKAQTHILERTPLTPGTEHREVRVOTGKEYEITKLNINRVDSMKITDG 547
    |||||||
OY 455 ----- 454
DB 548 AASTFDLINVVQRIGIELDNAGNVTKTETKIIAKLGEGDDNVFVGSCTTEIDGGEGYD 607
    |||||||
OY 455 ----- 454

```

```

DB 608 RVHYSRNGALTIIDATKETEQGSYTVNRFVETGKALHEVTSHTALVGNREEKIEYRHS 667
    |||||||
OY 455 ----- 454
DB 668 NNOHAGYTKDTLKAVEEIICTSHNDIFKSGKFNDAFNGGCVDTIIDGNDGNDRLFGK 727
    |||||||
OY 455 ----- 454
DB 728 GDDIIDGNGDDFIDGKGNDLHAGKGDDIFVHRKGDGNDIITDSGNDKLSFSDSNLK 787
    |||||||
OY 455 DLTFEYKHNLVITNSKKEKVTIQQNFREADFAKEVPNPKATDEKIEEIIIGONGERITS 514
    |||||||
DB 788 DLTFEYKHNLVITNSKKEKVTIQQNFREADFAKEVPNPKATDEKIEEIIIGONGERITS 847
    |||||||
OY 515 KOYDDLIANGNGKITQDELISKVVDNYELLKHSKNVTNSLDKLISVSAPFTSSNDSRNLV 574
    |||||||
DB 848 KOYDDLIANGNGKITQDELISKVVDNYELLKHSKNVTNSLDKLISVSAPFTSSNDSRNLV 907
    |||||||
OY 575 APTSMLDQSLSSIQFARGS 593
    |||||||
DB 908 APTSMLDQSLSSIQFARGS 926
    |||||||

```

```

Search completed: March 2, 2001, 10:54:48
Job time: 613 sec

```

OM of: US-09-306-689-12 to: Issued\_Patents\_AA:\* out\_format: pfs  
Date: Mar 2, 2001 10:52 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=framet\_n2p.model -DEV=xip  
-O/cgnt2\_1/USPTO.spool/US09306689/runat\_02032001\_102824\_9539/app\_query.fasta.1.2389  
-DB=Issued\_Patents\_AA -OPMT=fastan -SUFFIX=rai -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bl0sum62  
-TRANS=human40.ctd -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTMT=pfs  
-NORM=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US09306689 -CGCN1\_1\_98 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPXY  
-WAIT -THREADS=1

## Search information block:

Query: US-09-306-689-12  
Query length: 2088  
Database: Issued\_Patents\_AA:\*  
Database sequences: 174772  
Database length: 17957048  
Search time (sec): 62.480000

## score\_list:

Seq	Strd Orig	ZScore	E_Score	Len	Documentation
/cgnt2_6/p/ptodate/2/1aa/5A.COMB	pep:US-08-258-188-2	403.00	574.09	1.7e-25	758
/cgnt2_6/p/ptodate/2/1aa/5A.COMB	pep:US-08-526-813-2	403.00	574.09	1.7e-25	758
/cgnt2_6/p/ptodate/2/1aa/5A.COMB	pep:US-07-690-983D-47	368.00	545.39	5.9e-23	84
/cgnt2_6/p/ptodate/2/1aa/5A.COMB	pep:US-08-387-156-4	290.00	435.73	1.3e-16	49
/cgnt2_6/p/ptodate/2/1aa/5B.COMB	pep:US-08-694-865-4	290.00	435.73	1.3e-16	49
/cgnt2_6/p/ptodate/2/1aa/5B.COMB	pep:US-08-878-748-4	290.00	435.73	1.3e-16	49

seq\_name: /cgnt2\_6/p/ptodate/2/1aa/5A.COMB.pep:US-08-258-188-2

## seq\_documentation\_block:

Sequence 2, Application US/08258188  
Patent No. 5475098  
GENERAL INFORMATION:  
APPLICANT: HALL, Robert H.  
APPLICANT: XU, Jian Guo  
TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.  
TITLE OF INVENTION: coli 0157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND  
SPECIFIC DETECTION OF 0157:H7 AND OTHER ENTEROHEMORRHAGIC  
TITLE OF INVENTION: E. coli  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Steuart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
COMPUTER TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/258,188  
FILING DATE: 14-JUN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Garrett-Wackowski, Eugenia  
REGISTRATION NUMBER: 37,330

REFERENCE/DOCKET NUMBER: 15280206, DHSSEL35940  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 758 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-258-188-2

alignment\_scores:  
Quality: 403.00 Length: 395  
Ratio: 1.807 Gaps: 11  
Percent Similarity: 56.456 Percent Identity: 30.633

alignment\_block:  
US-09-306-689-12 x US-08-258-188-2

Align seg 1/1 to: US-08-258-188-2 from: 1 to: 758

964	GCACCTGACTGTCGACATATAAATGCTTCAACAGCTAAATAAGTGGTGC	1013
1	SerPheIleuGlYasnSerSapAlaHisThrGlyThrLysAlaIal	17
1014	GGGTTTGAATTGGCAACCAAGTTGGTAATATACCAACCGCTT	1063
17	aglyIleGluLeuThrThrGluValLeuGlyAsnValIglyLysAlaIal	34
1064	CTCTTACATTTAGCCCAACGCTGTCGACAGGTTATCTTAACTGGG	1113
34	erGlnIleuLeuAlaGlnArgMetAlaGlnIleuSerThrAla	50
1114	CCGTGCTGCTTAAATGCTTCTTACGTTCTTCCGATTAGCCATT	1163
51	AlaSerAlaGlyLeuIleThrSerAlaValMetLeuAlaIleSerProle	67
1164	AGCATTTGCCGCTATTGCGCATTAATTAATCATGCAAAAGTTAGAGA	1213
67	userPheLeuAlaAlaIalAspLysPheGluArgAlaLysGlnLeuGlu	84
1214	GTTATGCCGACGCTTTAAATAATAGCGTAGCAGGATTAATTATTA	1263
84	erTyrSerGluArgPheLysLysLeuAsnTyrGlnLysAlaLeuLeu	100
1264	GCAGATATCAGCGGGAACAGGACTATTGATCGCGTTACGTCAMP	1313
101	AlaAlaPheHisLysGluThrGlyAlaIleSapAlaLeuThrThrI	117
1314	TATACCGCATTTGCCGCTATTGCTGCTGCTGCTGCTGCTGACCC	1362
117	eaSnThrValLeuSerSerValSerAlaGlyAlaSerAlaIleSerSera	134
1362	.....	1362
134	laSerLeuIleGlyAlaProIleSerMetLeuValSerAlaLeuThrGly	150
1363	.....GATTAACTTTGAAAGT	1382
151	ThrIleSerGlyIleLeuGluAlaSerLysGlnAlaMetSerPheGluI	167
1383	TAAACATAATCTTGTGCATCAGCAATAGCAAAAGAAAGTACCATTC	1432
167	IlaGluLysPheAlaIalArgIleAsnGlnLysGlnLysGlnLysGlyL	184
1433	AAACGTGTTCCGAGAG	1449
184	ysAsnTyrPheGluAsnGlyTyrAspAlaArgHisAlaIalPheLeuGlu	200
1450	.....GCTGATTTTGCTAAAGAAGTCCGTAATATATA	1481

```

201 AspSerLeuSerLeuAlaAspPheSerArgLinhIsAlaValGluAr 217
1482 AGCAACTAATA.....GATGAGAAATCGAACAATCA 1513
      |||:|||||:|||||:|||||:|||||:|||||:
217 gAlaValAlaIleThrGlnGlnHisTrpAspGluLysIleGlyIuLeuA 234
1514 TCGGT.....CAAAATGGCGAGCGGATCACCTCAAGAAGATTGAT 1557
      |||:|||||:|||||:|||||:|||||:|||||:
234 laGlyIleThrArgAsnAlaAspArgSerGlnSer..... 245
1558 CTTATGCCAAAGGTAAAGCGCAAAATTACCAAGATGAGTATCAAAAGT 1607
      |||:|||||:|||||:|||||:|||||:|||||:
246 .....GlyLysAlaTyrIle.....AsnTyr 252
1608 TGTGATACATGATGATGCTCAAA...CATGCAAAATGTCACAACA 1654
      |||:|||||:|||||:|||||:|||||:|||||:
252 rLeuGlnAsnGlyGlyLeuLeuGlnAlaGlnProLysGlnPheThrGln 269
1655 GCTTA...GATTAAGTTAATCTCATCTGTAGTGCATTTACCTGCTAAT 1701
      |||:|||||:|||||:|||||:|||||:|||||:
269 InValPheAspProGlnLysGlyThrIleAspLeuSerThrGlyAsnVal 285
1702 GATTGCAAAATGATTAATAGTGCCTCAACTTCATGTTGATCAAAAGTTT 1751
      |||:|||||:|||||:|||||:|||||:|||||:
286 SerSerValIleThrPheIleThrProThr..... 295
1752 ATCTTCTCTCAATTTGCTAGGGGATCTCAGCATTTGAGCTAGCGCCGTC 1801
296 .....PheT 297
1802 GCCCTGGCAGCGGTTCTCAAGATTGAGTACGGCTGCTCCGGTGCC 1851
      |||:|||||:|||||:|||||:|||||:|||||:
297 hrProGlnGluGluValArgGlu..... 304
1852 TCTAGCCAGCATTTGAGCTAGCGCTGCGCCCTGCGACGGTACAGAGA 1901
305 .....ArgLysGlnSerGlyTyrG1 312
1902 TTGAGAGTACGCGCTGCGCTCGGGTGGATGTCAGCATTTGGAGCTAC... 1948
      |||:|||||:|||||:|||||:|||||:|||||:
312 uTyMetThrSerLeuIleValAsnGlyLysAspThrTTPSerValLysG 329
1949 GCGTCGCGCCTGGCAGCGGTTCTCAAGATTGAGC 1983
      |||:|||||:|||||:|||||:|||||:|||||:
329 LyIleLysAsnHisLysGlyValTyrAspTyrSer 340
seq_name: /cgn2.6/ptodata/2/1aa/5A_COMP.pep:US-08-526-813-2
seq_documentation_block:
; Sequence 2, Application US/08526813
; Patent No. 5756293
; GENERAL INFORMATION:
; APPLICANT: Hall, Robert H.
; APPLICANT: Xu, Jian Guo
; TITLE OF INVENTION: A New and Distinctive DNA Sequence of E.
; Patent No. 5756293
; TITLE OF INVENTION: coli 0157:H7 and its Use for the Rapid, Sensitive and
; TITLE OF INVENTION: Specific Detection of 0157:H7 and Other Enterohemorrhagic
; TITLE OF INVENTION: E. coli
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/526,813
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,188
; FILING DATE: 14-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 15280-206-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-526-813-2

alignment_scores:
      quality: 403.00      length: 395
      ratio: 1.807      gaps: 11
Percent Similarity: 56.456      Percent Identity: 30.633

alignment_block:
US-09-306-689-12 x US-08-526-813-2 ..

Align seq 1/1 to: US-08-526-813-2 from: 1 to: 758

964 GCACTGTACTGTGCGATAAAATGCTTCAACAGCTAAAGTGGGTGC 1013
      |||:|||||:|||||:|||||:|||||:|||||:
1 SerPheIleuGlnAsnSerAspAlaHisThrGlyThrLysAlaIleAl 17
1014 GGGTTTGAATTGGCAACACAGTTGTGTAATATTACCAAGCCGTTT 1063
      |||:|||||:|||||:|||||:|||||:|||||:
17 agLyIleGlnIleThrThrGlnValIleuGlnAsnValGlyLysAlaValS 34
1064 CTCTTTCATTTTACCCCAACGCTGTCACAGGTTATGTTCACTGCGG 1113
      |||:|||||:|||||:|||||:|||||:|||||:
34 erGlnTyrlleuAlaGlnArgMetAlaGlnGlyLeuSerThrThraLa 50
1114 CCTGCGCTGTTAATTGCTGCTACTGTTCTCTTCCGATTAGCCCAT 1163
      |||:|||||:|||||:|||||:|||||:|||||:
51 AlaSerIleGlyLeuIleThrSerAlaValMetLeuAlaIleSerProLe 67
1164 AGCAATTGCCGGTATTGCCGATAATTAAATTCATGCAAAAAGTTAGAGA 1213
      |||:|||||:|||||:|||||:|||||:|||||:
67 uSerPheLeuAlaIleAlaAspLysPheGluArgAlaLysGlnIleuGlnS 84
1214 GTTATGCCGAGCGTTTAAATAATTAGCTATGACGAGATTAATTATTA 1263
      |||:|||||:|||||:|||||:|||||:|||||:
84 erTySerGlnArgPheLysLysLeuAsnTyrgLysAspAlaLeuLeu 100
1264 GCAAAATATGAGCGGAGACAGCATTTGATGATGATGATGATGATGAT 1313
      |||:|||||:|||||:|||||:|||||:|||||:
101 AlaIlePheHisLysGlnThrGlyAlaIleAspAlaIleAlaLeuThrTrI 117
1314 TAATACCGCATTTGCCGCTATTGCTGTGTGTGTGTGTGTGTGTGTG 1362
      |||:|||||:|||||:|||||:|||||:|||||:
117 eaSnThrValLeuSerSerValSerAlaGlyValSerAlaIleSerSera 134
1362 ..... 1362
134 laSerIleuIleGlyAlaProIleSerMetLeuValSerAlaLeuThrGly 150
1363 .....GATTACATTGGAAGAAGT 1382
151 ThrIleSerGlyIleLeuGlnAlaSerLysGlnAlaMetPheGlnHisVa 167
1383 TAACATATCTTGTCTATCAGATACGAAACAAAGAGAAAGTGCATTC 1432

```

```

167 lalagluyspheaialaargileansglutrrgluylsgluhieglyl 184
1433 AAAATGCTCCGAGAG..... 1449
184 yASnltYrthegluansnglytYrAspAlaRghisAlaAlaPheleuglu 200
1450 .....GCTGATTTTCTAAAGAGTGCCTAATTATPAA 1481
201 AsPserleuSerleuLeuAlaAsPserArglnHisAlaValGluAr 217
1482 ACCAACTAAA.....GATGAGAAATTCGAAAGAAATCA 1513
217 galavalAlaIlethnglnGlnHisrAspArgluylsleglygluLeuA 234
1514 TCGGT.....CAAAATGCGGAGCGGATTCACCTCAAGCAAGTGTGAT 1557
234 lAglylYlethrArghsAlaAsPargSerGlnSer..... 245
1558 CTATTCGCAAAAGGTAAAGCAAAATTACCAGATGAGCTATCAAAAGT 1607
246 .....GlylYsAlaTyrIle.....AsnTyr 252
1608 TCTTGATTAATGATGATTCCTCAAA...CATAGCAAAATGTCACAACA 1654
252 rleugluansnglygluLeuGlnAlaGlnProlysgluPheThrglng 269
1655 GCTTA...GATTAAGTTAATTCATCATCTGTAAAGTCATTACCTCGTCTAA 1701
269 lvalPheAsPProGlnlysglyThrlleAsPleuSerThrglyAsnVal 285
1702 GATTGAGAAATGTATAGTGGTCCAACTTCATGTTGGATCAAAAGTTT 1751
286 SerSerValleuThrPheIleThrProthr..... 295
1752 ATCTTCTCTTCAATTGCTAGAGGATCTCAGCATTCGAGCTAGCGCTGC 1801
296 .....Phet 297
1802 GCCCTGGCAGCGGTTCTCAGATTGGAGCTAGCGCTCGCGGGTGC 1851
297 hPrrGlygluGlnValArglu..... 304
1852 TCTAGCAGCATTTGGAGCTAGCGGCTCGCGCTGGCAGCGGATGAGCA 1901
305 .....ArglysglnSerGlylystYrG 312
1902 TTGGAGCTAGCGCTCGCTCGCGGATGATCAGCATTTGAGCTAC...G 1948
312 utYrMetThrSerleuIleValAsnGlyAsPThrTrpSerValYsg 329
1949 GCCTGCGCCCTGGCAGCGGTTCTCAGATTGGAGC 1983
329 lyIlelyAsnHislysglyValYrAsPTrSer 340

seq_name: /cgn2_6/ptodata/2/1aa/PCUTUS_COMB.pep: PCT-US95-08554-2

seq_documentation_block:
; Sequence 2, Application PC/TUS9508554
; GENERAL INFORMATION:
; APPLICANT: HALL, Robert H.
; APPLICANT: XU, Jian Guo
; TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
; TITLE OF INVENTION: coli 0157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
; TITLE OF INVENTION: SPECIFIC DETECTION OF 0157:H7 AND OTHER ENTEROHEMORRHAGIC
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US

```

```

? ZIP: 94105-1493
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/08554
? FILING DATE: 14-JUN-1994
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Garrett-Wackowski, Eugenia
? REGISTRATION NUMBER: 37,330
? REFERENCE/DOCKET NUMBER: 15280206, DHHS135940
? TELEPHONE: (415) 543-9600
? TELEFAX: (415) 543-5043
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 758 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? PCT-US95-08554-2

Alignment_scores:
Quality: 403.00 Length: 395
Ratio: 1.807 Gaps: 11
Percent Similarity: 56.456 Percent Identity: 30.633

Alignment_block:
US-09-306-689-12 x PCT-US95-08554-2 ..
Align seg 1/1 to: PCT-US95-08554-2 from: 1 to: 758

964 GCACCTGTACTGTGACGATTAATAAGCTTCACAGCTAAATAAGTGGGTGC 1013
1 SerPheIleLeuGlyAsnSerAspAlaHisThrGlyThrlYsAlaAlaI 17
1014 GGGTTTGAATTGGCAACCAAGTTGTTGTAATATTACCAACCGCTT 1063
17 aglyIleGluLeuThrThrglnValleuGlyAsnValGlylsAlaValS 34
1064 CTTCCTTACATTTTAGCCCAAGCTGTTCAGCAGGATTATCTTCAACTGG 1113
34 erGlnIYrIleleuAlaglnArgMetAlaGlnGlyleuSerThrThrAla 50
1114 CCGTGGCGTCTTAATGCTCTACTGTTCTGCTGCGATTAGCCCAT 1163
51 AlAserAlaGlyleuIleThrSerAlaValMetLeuAlaIleSerProle 67
1164 ACATTTGCCGATTTGCCGATTAATTTAATCATGCAAAAGATTAGAGA 1213
67 userPheleuAlaAlaAlaAsPlySphelunrghAlalysGlnleugluS 84
1214 GTTAGCCGAGCGCTTAAAAATTAGCTATGACGAGCATATAATTATTA 1263
84 ertYrSerGlnrghPheIlySlySleuAsnTyrGlnGlyspAlaLeuLeu 100
1264 GCAGATATACGGGGGAGAGGAGCTATTGATGATCGGTACGTCACAT 1313
101 AlAlaAlaPheHislysgluThrGlyAlaIleAspAlaAlaLeuThrI 117
1314 TATACCGCATTTGCCGCTATTTGCTGGTGTGCTGCTGCTGCAGCC 1362
117 eAsnThrValleuSerSerValSerAlaIleValSerAlaIleSerSera 134
1362 ..... 1362
134 lAserleuIleGlyAlaProIleSerMetleuValSerAlaLeuThrgly 150
1363 .....GATTTACACTTGAAAAAGT 1382

```

```

151 ThrIleSerGlyIleuGluAlaSerLysGlnAlaMetPheGluHisVal 167
1383 TAACACATATCTGTGTCATCAAGATAGCAAAAAGAGAAAGTGCACATTC 1432
167 LAlaGluLysPheHisAlaAlaArgIleAsnGluTTPGluLysGluHisGlyL 184
1433 AAAACGTGTTCCGAGAG..... 1449
184 yAsnGlyPheGluAsnGlyLysAlaArgHisAlaAlaPheLysGlu 200
1450 .....GCTGATTTCCTAAGAACTGCTTAATTATTA 1481
201 AspSerLeuSerLeuAlaAspPheSerArgGlnHisAlaValGluArg 217
1482 AGCAACTAAA.....GATGAGAAATGCAAGAAATCA 1513
217 gAlaValAlaIleThrGlnGlnHisTTPAspGluLysIleGlyGluLeuA 234
1514 TCGGT.....CAAAATGGCGAGCGATCACCTCAAGCAAGTTCATGAT 1557
234 laGlyIleThrArgAsnAlaAspArgSerGlnSer..... 245
1558 CTATGCAAAAGGTACGCCAAATTAACCAAGATGATGCTATCAAAAGT 1607
246 .....GlyLysAlaTyrIle.....AsnTyr 252
1608 TGTTGATACATGATGATTCCTCAAA.....CATAGCAAAATGTCACAAACA 1654
252 rLeuGluAsnGlyGlyLeuLeuGluAlaGlnProLysGluPheThrGlnG 269
1655 GCCTTA..GATTAAGTTAATCTCATCTGTAAGTCATTAACCTGCTCAAT 1701
269 lValAlaPheAsnProGlnLysGlyThrIleAspLeuSerTnGlyAsnVal 285
1702 GATTCGAGAAATGATTATAGTGGCTCCAACTTCATATGTTGATCAAAATT 1751
286 SerSerValLeuThrPheIleThrProThr..... 295
1752 ATCTTCTCTCAATTTGCTAGGGATCTCAGCAATTTGAGCTAGCGCCTGC 1801
296 .....PheT 297
1802 GCCCTGGCAGCGGTTCTCAAGATTGAGTACGCGCTGCGGTCGCGGTCGC 1851
297 hrProGlyGlnGluValArgGlu..... 304
1852 TCTAGCCAGATTGGAGCTAGCGCCTGCGCTGGCAGCGGTAGCCACA 1901
305 .....ArgLysGlnSerGlyLysTyrGln 312
1902 TTGGAGCTAGCGCTGCGCTGCGGTGATGTCATGAGCATTTGAGCTAGC...G 1948
312 uTyrMetThrSerLeuIleValAsnGlyLysAspThrTTPSerValLysG 329
1949 GCCTGCGCCTGGCAGCGGTTCTCAAGATTGAGC 1983
329 lTyrLeuAsnHisLysGlyValTyrLysTyrSer 340
seq_name: /cgn2_6/plodata/2/1aa/5A_COMB.pep:US-07-690-983D-47
seq_documentation_block:
; Sequence 47, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: F. J. Lardner
; STREET: 3000 K Street, N.W.

```

```

; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-690-983D-47

```

```

alignment_scores:
  Quality: 368.00      Length: 101
  Ratio: 4.779        Gaps: 7
  Percent Similarity: 76.238  Percent Identity: 68.317

```

```

alignment_block:
US-09-306-689-12 x US-07-690-983D-47 ..

```

```

Align seg 1/1 to: US-07-690-983D-47 from: 1 to: 84

```

```

22 TCTGACGATTGGAGCTACGGCCTGCGCTGGCAGCGGTTCTCAAGTTG 71
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 SerGlnHisTTPSerTyrGlyLeuArgProGly.....GlnHisTTP 15
72 GAGCTAGCGCCTGCGCGGTTGAGCTAGCCAGATTGAGCTAGCGGTCGCGGT 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15 pSerTyrGlyLeuArgProGly.....GlnHisTTPSerTyrGlyL 29
122 TCGCCTGCGCAGCGGTAGCCAGATTGAGCTAGCGCCTGCGCTGCGGTT 171
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
29 euArgProGly.....GlnHisTTPSerTyrGlyLeuArgProGly 42
172 GAGCTAGCGATTGGAGCTAGCGCCTGCGCTGGCAGCGGTTCTCAAGA 221
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
43 .....GlnHisTTPSerTyrGlyLeuArgProGly.....GlnH 54
222 TTGGAGCTAGCGCCTGCGCTGCGGTTGAGCTAGCCAGATTGAGCTAGC 271
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
54 sTTPSerTyrGlyLeuArgProGly.....GlnHisTTPSerTyrG 68
272 GCCCTGCGCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCCTGCGCTCGC 321
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
68 lLeuArgProGly.....GlnHisTTPSerTyrGlyLeuArgPro 81
322 GGT 324
  |||
82 Gly 82

```

```

seq_name: /cgn2_6/plodata/2/1aa/5A_COMB.pep:US-08-387-156-4
seq_documentation_block:
; Sequence 4, Application US/08387156

```



```

; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-387-156-4

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-12 x US-08-387-156-4  ..
Align seg 1/1 to: US-08-387-156-4 from: 1 to: 49

25 CACGATGGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 17
rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34

75 CTACGGCCTCGTCGCGGTGGTCTTACCCAGCATTTGAGCTACGGCCTGC 124
|||||
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34

125 GCCCTGGCAGCGGTAGCCAAAGATTGGAGCTACGGCCTGCGCGGT 171
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-694-865-4
seq_documentation_block:
; Sequence 4, Application US/08694865
; Patent No. #5837268
; GENERAL INFORMATION:

```

```

; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-694-865-4

```

```

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-12 x US-08-694-865-4  ..

```

```

Align seg 1/1 to: US-08-694-865-4 from: 1 to: 49

25 CACGATGGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 17
rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34

75 CTACGGCCTCGTCGCGGTGGTCTTACCCAGCATTTGAGCTACGGCCTGC 124
|||||
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34

125 GCCCTGGCAGCGGTAGCCAAAGATTGGAGCTACGGCCTGCGCGGT 171
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-878-748-4
seq_documentation_block:
; Sequence 4, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO

```

STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/878,748  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-878-748-4

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-12 x US-08-878-748-4 ..

Align seg 1/1 to: US-08-878-748-4 from: 1 to: 49

```
25 CAGATTGAGCTACGCGCTGCGCTGCGCTGCTCAAGATTGAG 74
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnAspTrpSe 17
75 CTACGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 124
|||||
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34.
125 GCCCTGCGAGCGGTAGCCAGATTGAGCTACGCGCTGCGGCT 171
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49
```



```

Db 165 SVLGININNEALIISSGSELELAKAGVDLASELVGNIAKGTATIEAFSEQJONFQKLVQNA 224
Qy 287 KGLTGLDCKLNKNI--LDRKGLDGLDYISGLLSATALVADKNASTAKKVGAGFELAN 344
Db 225 KGLGVGGOOLIISSSALSKTGLDIISSLSGVTSFPLADKNASTSKTVAGFELSN 284
Qy 345 QVNGRTIAVASSYIIAQAAGLSTSTGPVALLASTVASIISPFLAFGIADKFNHAKSLE 400
Db 285 QVIGGRTIAVASSYIIAQLAGLSTTGPAALLASSLSLSPISFLRVDNFRKSDIR 344
Qy 405 SYAEFFKKLGVDGNLLAEYORGTTIDASTATINTALAAIAGVSSAAADL----- 456
Db 345 EFAEFKKLTGVBGKLLSDFLFHAEGTIDASTITSTLSATAACTAASGALVGARTL 400
Qy 457 -----TFEKVHNIVITNSKKEKVTIONMFE----- 483
Db 405 LVYGTITGLISGLEFSKQPMLEHVASKLGTKIEEMERKYGNPENGVDARHKAFLEDSL 464
Qy 484 ---ADFAEVENYKATK-----DEKTEFELIGONGERTSKQVNDLIIAKNGKTTODE-- 532
Db 465 SLLSEFNQYETERAVLITQORMDEYIGELAGVYTK-----GDKISGKAAVVDYPEEG 517
Qy 533 --LSKVVNYE-----LLKSKNVNTLIDKLISSVAFSTSNDR 570
Db 518 KILAKKPPDENRVILIDPKKKGIIDSNGTSTILKFTVPL--LTPGTSR 564

```

RESULT	2	
09LC58		
ID	09LC58	PRELIMINARY; PRT; 998 AA.
AC	09LC58;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)	
DE	HEMOLYSIN A.	
GN	EHCC-HLYA.	
OS	Escherichia coli.	
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;	
OC	Escherichia.	
OX	NCBI_Taxid=562;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Tanabe I., Wakisaka-Saito N., Harada Y., Zhang H.-M., Yamamoto T.;	
RT	"The enterohemorrhagic Escherichia coli (EHEC)-hemolysin genes of a	
RT	Shiga toxin 1 (Stx1)- and Stx2-producing, serotype O128 Escherichia	
RT	coli strain with a greatest hemolytic activity.";	
RL	Acta Med. Biol. 0:0-0(2000).	
DR	EMBL, AB032930; BAA93708.1; -	
SO	SEQUENCE 998 AA; 107196 MW; CD7A8BE9D62D56 CRC64;	

Query Match	23.8%	Score 855	DR 2	length 998
Best Local Similarity	35.8%	Pred. No. 1e-36		
Matches 220; Conservative	92;	Mismatches 182;	Indels 120;	Gaps 15;

```

0Y 110 S$PPTGK$K$IT$Y$P$W$Y$D$T$E$O$G$N$D$Y$A$A$E$B$E$J$E$O$R$E$R$R$N$I$A$T$A$O$S$T$G 169
Db 25 S$S$V$A$S$A$C$K$T$U$L$L$P$D$N$Y$E$---$A$Q$G$V$I$N$T$Y$A$A$D$D$G$E$H$H$R$B$R$D$T$A$N$O$E$F$G 81
0Y 170 I$Q$T$A$G$T$E$R$O$V$I$S$A$P$O$D$K$U$L$O$K$-T$K$A$G$A$L$-G$S$E$S$T$Y$O$N$A$K$K$Y$U$S$G$I$O$S$T$U$G$ 227
Db 82 A$E$K$V$G$T$E$R$E$V$A$F$A$P$O$D$U$L$K$Y$O$K$V$G$K$U$G$T$E$A$E$N$G$N$N$G$K$A$G$Y$U$S$A$O$N$T$G$I 141
0Y 228 V$A$G$N$D$D$E$A$O$-----N$N$S$O$H$A$R$A$G$E$T$N$S$L$E$N$T$A$N$S$Y$K$T$U$D$E$G$E$O$S$O$F$G$K 282
Db 142 A$G$S$A$D$D$E$L$R$K$O$R$E$G$E$D$I$O$N$D$I$A$K$S$T$E$L$I$N$Q$U$D$V$Y$S$T$S$T$D$S$E$O$U$N$O$G$S$F 201
0Y 283 L$Q$N$K$G$T$E$D$D$K$N$G$G$U$D$K$A$G$E$D$V$I$S$G$L$S$G$A$T$A$U$V$I$A$D$K$A$S$T$A$K$V$G$A$G$E$L 342
Db 202 L$S$K$R$P$L$S$S$V$G$K$Q$N$U$P$N$G$P$D$G$E$D$V$Y$G$S$I$S$A$V$S$A$S$T$I$G$N$S$D$A$H$G$T$A$A$G$E$L 261
0Y 343 A$N$O$V$G$N$T$K$A$V$S$Y$U$A$O$R$V$A$A$G$S$T$G$P$V$A$U$U$A$S$T$V$A$S$U$S$A$P$L$A$F$A$G$I$D$K$F$N$H$A$K$ 4020

```

```

Db      262  TTQVIGNVKRAVSQYLLAQRMAQIGSTTAASAGLITSVAMHLISPLSEFLIADKFEBAKQ 321
QY      403  LESYAEKKKLGIDGDNLLAETQKGTIDASVTAINTALAIAGVSAABA----- 454
Db      322  LESYSEERKKLNTEGDALLAFHKEETADIALTTINTVLVSAGVSAASSASLIGAPI 381
QY      455  -----DLTEPKVHNLTVTSKKEXTIQNWFPE----- 483
Db      382  SMLVSALTGTISGLLEASKQAMFEHVADKFAARINEMKEHGKRYFENGVDARHAAFLD 441
QY      484  -----ADFAKEVPYKATK-----DEKIEIIL--ONGERITSQVDLLIAGNGKITQ 530
Db      442  SLSLADSTRQHNABRAVAITIQOHMDKIGLACITNABRSQ-----GKAYI-- 490
QY      531  DELSKVVDNYELLK-HSKNVTNSL-DKLISSVAFSTSSNDNRVLAPTSMLDOSLSLQ 588
Db      491  ---NYLENGILLAEQPREFIQYFDPQKGIIDLSTGNVSSVLTFFVPT----- 535
QY      588  FARGSQHWSYGLRPGSGQSDMSYGLRPGSGSQHWSYGLRPGSGSDMSYGLRPGSGQHS 648
Db      536  -----FTPGEEVRE-----RKQSGKEYEYMTSLVINGKDTWS 566
QY      649  Y-GLRPGSGSDMS 661
Db      567  VKGIRNKKGYTDYS 580

```

RESULT	3		
OC85101			
ID	085101	PRELIMINARY;	PRT; 998 AA.
AC	085101;		
DT	01-NOV-1998 (TReMBLrel. 08, Created)		
DI	01-NOV-1998 (TReMBLrel. 08, last sequence update)		
DR	01-JUN-2000 (TReMBLrel. 14, last annotation update)		
DE	HEMOLYSIN.		
GN	EHXA.		
OS	Escherichia coli.		
OG	Plasmid EHEC-hemolysin.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Escherichia.		
OX	NCBI_TaxId=562;		
RV	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ENTEROHEMORRHAGIC EC920006;		
RA	MEDLINE=98261495; PubMed=9596716;		
RX	Boerlin P., Chen S., Colbourne J.K., Johnson R., De Grandis S.,		
RA	Gyles C.;		
RT	"Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids		
FT	and the locus for enterocyte effacement in shiga toxin-producing E.		
RL	coli.";		
RI	Infect. Immun. 66:2553-2561(1998).		
DR	EMBL; AF043471; AAC24352.1; -.		
DR	INTERPRO; IPR001343; -.		
DR	PFAM; PF00353; hemolysincabiind; 2.		
DR	PRINTS; PR00313; CABDNCRPT.		
DR	PROSITE; PS00330; HEMOLYSIN-CALCIUM; 2.		
KW	plasmid.		
SO	SEQUENCE	998 AA; 107095 MW; ESA566B239DCBE55 CRC64;	

	Score	DB 2	Length	998
Query Match	23.6%			
Best Local Similarity	35.7%			
Matches 219, Conservative	92			
Mismatches 183, Indels	120			
Gaps	15			

	Query Match	23.6%;	Score 849;	DB 2;	Length 998;
	Best Local Similarity	35.7%;	Prod. No.2.1e-36;		
	Matches 219;	Conservative 92;	Indels 120;	Gaps	
Qy	110	SSPFTGAKKILIIIPONYOYDTFGSGGLDIDVKAAEELGIEVQREERNRIATAQISLGT	169		
		:::	:::	:::	:::
Db	25	SSSVASACKKILIIIPDNYE---AQGVGIMELVKAADLEIETIHRERDOTALANQPFGT	81		
Qy	170	IQTATIGTERGIVTSAPDIDKLIK--TKACQAL--GSAEIVSONANKATVLLSGISILGS	227		
		:::     :::	:::     :::	:::     :::	:::     :::
Db	82	AEKVGGLTEREVALFAFDQIDKLKLYOKVQSKIGCTATENGNNLGRKAGVYLSALQNFGTGI	141		
Qy	228	VLAGMDLDEALQ-----NNSNQHALAKAGELTNSLTENIANSVKTLDDEFGDQISQFGSK	282		

```

Db 142 ALSGMALDELLRKOREGDISQNDIAKSSIIEIINQVLVDVSSINSTVDSFSQNLQCSF 201
Qy 283 LONIKGLCTLDKLNIGGLDCKAGIDLVISGLSGAFAALVADKNASTAKKAGFEL 342
Db 202 LSSKRRLSSVGKQNLNPLDPLGDLGVSGILSAVSASFILGNSDHTGKAAGIEL 261
Qy 343 ANOVGNITKAVSSYIIAQRVAAGISSTGPAVALIATVSLAISPLAFAGIADKFNHKS 402
Db 262 TTQVIGNVKAVSQYIIAQRAAGLSTTAASAGLITSAVMLAISPLSLAIDKFERAKQ 321
Qy 403 LESYAEERKKGIDYDGNLLAEYQRTGTIDASVTAINTALAIAGVAAAA----- 454
Db 322 LESYERKKNYEDDALLAFHKESGAIIDALTTINTVLSVSGVSAASASLIGAPI 381
Qy 455 -----DLTFEKVKHNLVITNSKKKRVITQNMFR----- 483
Db 382 SMLVSALTGTISGLEASKQAMFEHVADKFAARINEMEKKNFYENGADARHAAFLSD 441
Qy 484 -----ADFAKEVPNYKATK-----DEKIEEIIIG--QNGERTISQOVDDIAKNGKITQ 530
Db 442 SLSLADFSRQHAAVRAVAITQOHMDEKIGELAGITRRADRQS-----GKAYI-- 490
Qy 531 DELSKVVDNYELLK-HSKNVTNSL-DKLISVSATSSNDSRNVLVAPTSMLDQSLSLQ 588
Db 491 -----NYLENGGLLEAQREFEQVDFPOKGTIDLTSGVSSVLTFITPT----- 535
Qy 589 FARGSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWS 648
Db 536 -----FTPEEVEVRE-----RKQSGKREYMTSLIVNGKDTWS 566
Qy 649 Y-GLRPGSSQDMS 661
Db 567 VKGIKNHKGVYDYS 580

RESULT 4
046716 PRELIMINARY: PRT: 998 AA.
AC 046716;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE HEMOLYSIN A (HLXA).
GN HLXA OR EHEC-HLXA.
OS Escherichia coli.
OC Plasmid p0157.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-0157:H7;
RX MEDLINE=98290540; PubMed=9628576;
RA Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,
RA Yatsudo H.C., Kubota Y., Yamachi T., Iida T., Yamamoto K., Honda T.,
RA Han C., Ohtsubo A., Kasamatsu M., Hayashi T., Kuhara S., Shinaawa H.;
RT "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
RT enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
RT outbreak."
RL DNA Res. 5:1-9(1998).
RN [2]
RN SEQUENCE OF 241-998 FROM N.A.
RC STRAIN-EHEC;
RA Hall R.H., Xu J., Walderhaug M.O.;
RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN-EDL933;
RX MEDLINE=98391744; PubMed=9722640;
RA Burland V., Shao Y., Perna N.T., Plunkett G., Sofia H.J.,
RA Blattner F.R.;
RT "The complete DNA sequence and analysis of the large virulence plasmid

```

```

RT of Escherichia coli O157:H7."
RL Nucleic Acids Res. 26:4196-4204(1998).
DR EMBL; AB011549; BAA31774.1; -.
DR EMBL; U12572; AAA20544.1; -.
DR EMBL; AF074613; AAC70116.1; -.
DR INTERPRO; IPR001343; -.
DR PFAM; PF00353; hemolysincbind; 2.
DR PRINTS; PR00313; CANNONGREP.
DR PROSITE; PS00330; HEMOLYSIN_CACTUM; 2.
RW Plasmid.
SQ SEQUENCE 998 AA; 107047 MW; BB3C0D36FCBB9EBD CRC64;

Query Match 23.68; Score 847; DB 2; Length 998;
Best Local Similarity 35.58; Pred. No. 2,7e-36;
Matches 218; Conservative 92; Mismatches 184; Indels 120; Gaps 15;

Qy 110 SSFPTGAKKIIIVPOVYOYDTEGNGLODVKAELGIEVOREERNIATAQTSIGT 169
Db 25 SSSVRSACKKILILIPDNYE--AOGVGINELVRADELGLIEIHTERDDTAIANQFFGA 81
Qy 170 IOTAIGTERGIVLSAPOIDKLQK-TRAGQAL-GSASEIVQNNANKAKTVLSGIQSLGS 227
Db 82 AEKVGLTERGVAIAPQIDKLQKYQKVGSKIGSTAEVNGNNGKACTVLSALQNFQGI 141
Qy 228 VLACGDDLEALQ-----NNSQHALAKGELTNSLNIANSVKTLDPEGEQISQSGSK 282
Db 142 ALSGMALDELLRKOREGDISQNDIAKSSIIEIINQVLVDVSSINSTVDSFSQNLQCSF 201
Qy 283 LONIKGLCTLDKLNIGGLDCKAGIDLVISGLSGAFAALVADKNASTAKKAGFEL 342
Db 202 LSSKRRLSSVGKQNLNPLDPLGDLGVSGILSAVSASFILGNSDHTGKAAGIEL 261
Qy 403 ANOVGNITKAVSSYIIAQRVAAGISSTGPAVALIATVSLAISPLAFAGIADKFNHKS 402
Db 262 TTQVIGNVKAVSQYIIAQRAAGLSTTAASAGLITSAVMLAISPLSLAIDKFERAKQ 321
Qy 403 LESYAEERKKGIDYDGNLLAEYQRTGTIDASVTAINTALAIAGVAAAA----- 454
Db 322 LESYERKKNYEDDALLAFHKESGAIIDALTTINTVLSVSGVSAASASLIGAPI 381
Qy 455 -----DLTFEKVKHNLVITNSKKKRVITQNMFR----- 483
Db 382 SMLVSALTGTISGLEASKQAMFEHVADKFAARINEMEKKNFYENGADARHAAFLSD 441
Qy 484 -----ADFAKEVPNYKATK-----DEKIEEIIIG--QNGERTISQOVDDIAKNGKITQ 530
Db 442 SLSLADFSRQHAAVRAVAITQOHMDEKIGELAGITRRADRQS-----GKAYI-- 490
Qy 531 DELSKVVDNYELLK-HSKNVTNSL-DKLISVSATSSNDSRNVLVAPTSMLDQSLSLQ 588
Db 491 -----NYLENGGLLEAQREFEQVDFPOKGTIDLTSGVSSVLTFITPT----- 535
Qy 589 FARGSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWS 648
Db 536 -----FTPEEVEVRE-----RKQSGKREYMTSLIVNGKDTWS 566
Qy 649 Y-GLRPGSSQDMS 661
Db 567 VKGIKNHKGVYDYS 580

RESULT 5
P71223 PRELIMINARY: PRT: 998 AA.
ID P71223;
AC P71223;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE EHEC-HEMOLYSIN.
GN EHEC-HLXA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```

OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=78-92;  
RX MEDLINE=97034824; PubMed=8880480;  
RA Schmidt H., Karch H.;  
RT "Enterohemolytic colic O111 strains from patients with diarrhea and  
RT hemolytic-uremic syndrome.";  
RL J. Clin. Microbiol. 34:2364-2367(1996).  
DR EMBL; X94129; CA63849.1; -  
DR INTERPRO; IPR001343; -  
DR PFM; PF00353; hemolysinCabd; 2.  
DR PRINTS; PR00313; CABNDNGRPT.  
DR PROSITE; PS00330; HEMOLYSIN\_CALCITUM; 2  
SO SEQUENCE 998 AA; 107255 MW; DA3EF078C7E4131E CRC64;

Query Match	23.5%;	Score 845;	DB 2;	Length 998;
Best Local Similarity	35.5%;	Pred. NO. 3.4e-36;		
Matches 218; Conservative	92;	Mismatches 184;	Indels 120;	Gaps 15;

QY	110	SSPFTGAKKIIITVPOKYOXTDQGNLODYAAEELGVEYRERNNATFQOTSLGT	169
Db	25	SSSVRSACKKIILLIPDNYE---AQGVCINELVAAABELGIEIHRTERDDTAIANOFEGA	81
QY	170	TOTALGTERGIVSAPOIDKLLQK-TKAGALG-SAESIVONANKATVLSGQSILGS	227
Db	82	AEKVVYGERVAIFAPOLDLQKQYGVSGSKIRTEBMVGNNGKKGVTJLSALONFTGI	141
QY	228	VIAGDDUDEALO-----NNSNMHLAKRAGLELTSLIENTANSVKTLDEGEQISQRFK	282
Db	142	ALSGADELIRKOREGEDDISQNDIANSSTELIQLQVDTVSSINSTVDSSEQUNQJSGF	201
QY	283	LQNTKGLTGLDKLKNIGLIDKAGLGLDVISLGLSGATTAALVLDADKNASTAKKVGAFEL	342
Db	202	LSKRRLSSVGGKLLQNPEDDGLSDGLDVLVSGILISAVASPTILGNSDAHGTKAAAGTEL	261
QY	343	ANQVGNITTKAVSSYIIIAQRYAAGLSTGPAAIIASTVSLAISPLAPAGIADKFNHAKS	402
Db	262	TTOVYGNNGKAVSOYIIIAORNAOGLSTTASAGLITSAMVLATISPLFLAADKFERAKO	321
QY	403	LESVAERPKKGYODNNLAEYORGTGTIDASVAINTALAIAGVSAAAA-----	454
Db	322	LESISERKKLYNEBDALLAFHKETGAIDALITITIVLSSVAGVSAASSASLIGAPI	381
QY	455	-----DLTFEKVXHNILVITNSKREKVTIQMWFE-----	483
Db	382	SMLVSAITGTISGLLEASKQAMFEHVAEKFARINEMEKHEGKNYFENGYDARHAAFLD	441
QY	484	-----ADRAKEVPNNKARK-----DEKIEIILG--ONGEITNSKOYDDLIANGNGKITQ	530
Db	442	SLSLSIADSRQHAVERAAVATIQOHMDEKITIGELATITRNADSSQ-----GKAVI--	490
QY	531	DELSKVYDNNYELTK-HSKNVTNSL-DKLISVSASFSTSSNDNRNLVAPTSMLODSSLQ	588
Db	491	-----NYLENGLLLEAQPREFQYFDDPKGITIDLSTGVSSVLFITPT-----	535
QY	589	FARQSQHWSTYGLRPGSGQDMSYGLRPGSGQSHWSTYGLRPGSGQDMSYGLRPGSGQHS	648
Db	536	-----FTTQGEVEYR-----RKQSGKEYMTFELVINGKDTWS	566
QY	649	Y-GLRPGSGQDMS	661
Db	567	VKGJKNHKGYTDYS	580
RESULT	6		
ID	047461		
MC	047461	PRELIMINARY;	PRT; 998 AA.
JT	01-NOV-1996 (TREMBLREL 01, Created)		

DT 01-NOV-1996 (TrEMBLrel. 11, last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 04, last annotation update)  
DE PLASMID-DNA FOR EHEC-HEMOLYSIN OPERON.  
GN EHEC-HLYA.  
OS Escherichia coli.  
OC Plasmid p0157.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_Taxid=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EDL 933;  
RX MEDLINE=95172699; PubMed=7868227;  
RA Schmidt H., Beutin L., Karch H.;  
RT "Molecular analysis of the plasmid-encoded hemolysin of Escherichia  
col. O157:H7 strain EDL 933".  
RL Infect. Immun. 63:1055-1061(1995).  
RL EMBL; X86087; CAA60042.1; -.  
DR INTERPRO: IPR001343; -.  
DR PFAM: PF00353; hemolysincabind; 2.  
DR PRINTS: PR00313; CABNDNGRPT.  
DR PROSITE: PS00330; HEMOLYSIN\_CALCIIIM; 2.  
KW plasmid.  
SQ SEQUENCE 998 AA; 107032 MW; 4DBBE108C309BF7E CRC64;

Query Match	23.5%	Score 845;	DB 2;	Length 998;
Best Local Similarly	35.5%	Pred. No. 3.4e-36;		
Matches 218;	Conservative 91;	Mismatches 185;	Indels 120;	Gaps 15;

QY	110	SSPRTKQKIIITVYPOVQYDTEGNOGLDQVAAEELIEVQREERRNIATPQTSLGT	169
Dd	25	SSSVRSACKKILLILLPDNYE---AQGYNELVYAADELIEIHRTERDOTALANOPFGA	81
QY	170	IQTAIGITERCIVLAPQIDKLLQK-TRGAOL-GSAESIYQANAKKATVLSGTOSILGS	227
Dd	82	AEKVYGLTERGVAFAPQDLDKLQKYKVGSKIGCTEENNGNCKAGVIALSALONFTGI	1414
QY	228	VLAGDDIDDEAO-----NNSNMHAAKAGLEMTSLIENIANSVKTLTDEGEOISOPGSK	282
Dd	142	ALSGADELILRKOREGEDISQNDIANSSTELIQLQVDTYSINSTYDSFSEQNLQJSTF	201
QY	283	LQNKIGLGLTLDKJKNIGGLDKAGLGLDVISIGLSLGATPAVLVLDKNA5PAKKVAGFEL	342
Dd	202	LSMKPRLLSSVGKGLONLPDLGJLGDGLDGYVSGILISAVSAPRLGNDAPHGTMAAGIEL	261
QY	343	ANOVVGNITKAVSSYTIILAQRYAAGLSTGPYAAIIASTVLSLAISPLAPAGIADKFNHAKS	402
Dd	262	TYQVYGNKGAIVSOYITIIAORAAOLSTTASASAGLITSAMVALISPLSFLLAADKFERAKO	321
QY	403	LESVAERPKKIKGYODNLIAFYOGTGTIDASVAINMTALAIAGVSAANA-----	454
Dd	322	LESTSERKKLNTGSDALLAGFHFKETGALIDAGLTTINTVLSSVAGVSAASSASLIGAPT	381
QY	455	-----DLTFEKVXHNILVITNSKKEKVTIONMFE-----	483
Dd	382	SMLV5ALGTIGTISGLEASKQMFHEVAEKFARINEMEKHEGKNYFENGYDARHAAFLD	441
QY	484	-----ADPAKEVPYKATK-----DEKIEBITG-ONGERTITSKOVDLLIAKGNKTI	530
Dd	442	SLSLLEPFSRCHAVERAIVAIITHQMDKEIGELAGITNNADRSGS-----GKAYI--	480
QY	531	DELSKVYDNYEELK-HSKNVTNSL-DKLIS5SAFTSSSNDNRNVLVAP7SMLODLSLQ	588
Dd	491	-----NYLENGLLAQPREFLTQYVFDQKGIIDLTSGNVSVLFTFT-----	553
QY	589	FARGSQHWSYGLRPGSGSQDMSYGLRPGSSQHWSYGLRPGSGSQDMSYGLRPGSQHWS	648
Dd	536	-----FTPGEEVRE-----RKQSGKYEYMTSLIVNGKDTMS	566
QY	649	X-GLRPGSGSQDMS 661	
Dd	567	VKGIKRHHKGVYDYS 580	







```

DT 01-AUG-1998 (TREMBLrel. 07, Created)
DR 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ALPHA HEMOLYSIN (FRAGMENT).
GN HYL.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR 43;
RX MEDLINE=98155142; PubMed=9495754;
RA Boyd E.F., Hartl D.L.;
RT "Chromosomal regions specific to pathogenic isolates of Escherichia
RT coli have a phylogenetically clustered distribution."
RL J. Bacteriol. 180:1159-1165(1998).
DR EMBL; AF037573; AAC38262.1; -
DR INTERPRO: IPR000911; -
DR PROSITE: PS00359; RIBOSOMAL_L11; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 181 AA; 19101 MW; A3586DC663B571E CRC64;

Query Match 10.5%; Score 377.5; DB 2; Length 181;
Best Local Similarity 45.9%; Pred. No. 6.8e-13;
Matches 83; Conservative 33; Mismatches 58; Indels 7; Gaps 3;

QY 135 GNGIADLVKAAEELIEVOERERNINATQTSIGTITGTERGIVLSAPQIDKLQK 194
DB 1 GSSINDVLTADDELIEVOYDERKNGTATKQVFGTAERKILGTERGVATFAFQDLKLQK 60

QY 195 -TKAGQAL-GSAESIYVONANKAKTVLSGIQSLGSLVAGMDDEALQ-----NNSQNAL 247
DB 61 YOKAGNKKLGSAENIGDNKAGSVLSTFQNFGLTALSSMKIDELIKOKSGSNVSSSEL 120

QY 248 AKAGLELTNSLENTANSVKTDEFEQISQFGSKLQNKGTIGDKLNKGIDKAGL 307
DB 121 AKASIELINQLVDTAASINNANNVNSFSQDLNKLGSVLSNTKHLNGVGNKLNLPNDNIGA 180

QY 308 G 308
DB 181 G 181

RESULT 13
ID 070070 PRELIMINARY: PRT; 181 AA.
AC 070070;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ALPHA HEMOLYSIN (FRAGMENT).
GN HYL.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR 48, ECOL 51, ECOL 52, ECOL 54, ECOL 60, ECOL 63;
RX MEDLINE=98155142; PubMed=9495754;
RA Boyd E.F., Hartl D.L.;
RT "Chromosomal regions specific to pathogenic isolates of Escherichia
RT coli have a phylogenetically clustered distribution."
RL J. Bacteriol. 180:1159-1165(1998).
DR EMBL; AF037579; AAC38263.1; -
DR EMBL; AF037575; AAC38264.1; -
DR EMBL; AF037576; AAC38265.1; -
DR EMBL; AF037577; AAC38266.1; -
DR EMBL; AF037578; AAC38267.1; -
DR INTERPRO: IPR000255; -

```

```

DR INTERPRO: IPR000911; -
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
DR PROSITE: PS00359; RIBOSOMAL_L11; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 181 AA; 19104 MW; 61716DD676D6D613 CRC64;

Query Match 10.4%; Score 372.5; DB 2; Length 181;
Best Local Similarity 45.3%; Pred. No. 1.2e-12;
Matches 82; Conservative 34; Mismatches 58; Indels 7; Gaps 3;

QY 135 GNGIADLVKAAEELIEVOERERNINATQTSIGTITGTERGIVLSAPQIDKLQK 194
DB 1 GSSINDVLTADDELIEVOYDERKNGTATKQVFGTAERKILGTERGVATFAFQDLKLQK 60

QY 195 -TKAGQAL-GSAESIYVONANKAKTVLSGIQSLGSLVAGMDDEALQ-----NNSQNAL 247
DB 61 YOKAGNKKLGSAENIGDNKAGSVLSTFQNFGLTALSSMKIDELIKOKSGSNVSSSEL 120

QY 248 AKAGLELTNSLENTANSVKTDEFEQISQFGSKLQNKGTIGDKLNKGIDKAGL 307
DB 121 AKASIELINQLVDTAASINNANNVNSFSQDLNKLGSVLSNTKHLNGVGNKLNLPNDNIGA 180

QY 308 G 308
DB 181 G 181

RESULT 14
ID 090617 PRELIMINARY: PRT; 1729 AA.
AC 090617;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PERICARDINE.
GN PRC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON R;
RA Chatter A., Astler M., Zaffran S., Semeriva M., Gratecos D.;
RT "Characterization of an extracellular matrix component specifically
RT expressed in the basal lamina of the dorsal vessel in Drosophila
RT melanogaster."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF203342; AAF13280.1; -
SQ SEQUENCE 1729 AA; 164661 MW; DA9B1B7FAB5CDEDB CRC64;

Query Match 5.8%; Score 208.5; DB 5; Length 1729;
Best Local Similarity 22.2%; Pred. No. 0.014;
Matches 185; Conservative 72; Mismatches 271; Indels 307; Gaps 45;

QY 13 YGLRPGSGSQDMS---YGLRPG-----GSSQHWSTYGLRPGSGSQDMS---YGLRPG-- 57
DB 417 YGTGPGIGGTGAGGPGYGTGPGIGAGGTGAGGPGYGTGPGIGGAGGPGYGTGPGIG 475

QY 58 ---GSOHWSYGLRPGSGSQDMS---YGLRPG-----GSSQHWSTYGLRPGSGSQDMS--- 102
DB 476 YGTGAGGPGYGTGAGGPGYGTGPGIGGAGGPGYGTGAGGPGYGTGAGGPGYGTGAGG 534

QY 103 -YGLRPG-----GSSFP-----KTGAKKILITIPONQYDTEGNGIADLVKAAE 146
DB 533 GTGTGPGIGAGGPGYGTGAGGPGYGTGAGGPGYGTGAGGPGYGTGAGGPGYGTGAGG 583

QY 147 ELGIEVOERERNINATQTSIGTITGTERGIVLSAPQIDKLQKTRKAGQ-ALG 202
DB 584 QPGYGTGPGVAGQGTGAGGPGYGTGAGGPGYGTGAGGPGYGTGAGGPGYGTGAGGPGYGT 639

```

QY	203	SAESIYQANAKKT---	VLSSIGSILISVLAGMDLDAIOMNSNQHLAAAGLELINSLI	253							
Db	640	SOPGIGGOTGAAGP	PGYGTQPPGVGAQGTGTGGPVGAGAPGT---	GGGTGAGGP-----	688						
QY	260	ENIANSVTLDEFEE	SOISOFESKIDONTKGLGT---	LDKLT---	KNIGGLDKAG	306					
Db	669	-----	YGRPPRIGGQTBP	PGYGTQPPGVGAQGTGTGGPVGAGAPGTGIGGOSGAG	736						
QY	307	L-----	-----	GIDVIGLSGATP	ALVALADKNASTAKKVGAGFELANQVGNIT	351					
Db	737	OPGVGSGP	GIGGQTGGGPGVGSIGGQTGA-----	GQPSYGSGQPGVGAONGGAGGP	PGYGTG	791					
QY	352	KAV-----	-----	SYTLADRVAA	GLSSTGPVAAIISTVSLAISLAPLA	391					
Db	792	RPVIGGGOTGAG	OPGVGGGTGVGSGSPGFLTQPGIGIS-	GPIGGKVGSGGSEAAKPGYVA	849						
QY	392	-----	-----	GIADKFNHAKSL	ESYMERFKLLYYDDDNLLAEYORGTGT---	DA	433				
Db	850	OPGIGGSPRYS	OPGIGDQTGAGS-----	-----	GYGGOPGISC	OTGGGPPGTGGQA	897				
QY	434	SVTAIN-----	-----	TALAAIAGVS	SAAAADLTFEKVKNHNLVITNSKKEKTYI	ONWPREAD	485				
Db	898	TISGLPGYGT	PGTGGALTAVPBG-----	HYGYE-----	-----	TPGIGGQTGTNNP	940				
QY	486	FAKEVPYNTKAT	KEKIEIITIGONCERIT	SKOVDLLAKNGKKT	TODELSKVVDNYELKH	545					
Db	941	FGGQ-----	-----	PGIGGOTGAG-----	OPGYGFLTQPGIG-----	968					
QY	546	SKNTVNSLDK	LISVSATTSNDRNLV	VAPTSMIDSL	SSLQPFARCSQMSYGLRPGSG	605					
Db	969	-----	-----	-----	OTGYSRKP	PGYGTQPGIG	986				
QY	606	SODMS-----	YGLRPG-----	SSSOHMSYGLRPGSGSO-----	DMSYGLRPG-----	GSQHW	647				
Db	987	GQTAAGQPGY	SGQPGIGGQTGAGP-	GTGSGTGTG	GGGILGAGQPGYSGP	PDIGGQTGAGP	104				
QY	648	SYGLRPGSG	SODMSYGLRPG-----	SSSOHMSYGLRPGSGSODMS-----	YGLRPG	693					
Db	1046	GYGAGP	GFHGGP-GYGNQPGVG	GGTQTAGP-GYSGQPGVG	GGTQTAGP	PGYGVIPG	1098				
RESULT 15											
Q9L469											
ID	Q9L469	PRELIMINARY;		PRT;		1706 AA.					
AC	Q9L469:										
DT	01-OCT-2000	(TReMBLrel. 15, Created)									
DT	01-OCT-2000	(TReMBLrel. 15, Last sequence update)									
DT	01-OCT-2000	(TReMBLrel. 15, Last annotation update)									
DE	BIUNCTIONAL	HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (EC 4.6.1.1).									
GN	CVAA.										
OS	Bordetella parapertussis.										
OC	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;										
OC	Bordetella.										
OX	NCBI_taxID=519;										
RP	[1]										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=63.2;										
RA	Boursaux-Bude C., Guiso N.;										
RT	"Sequencing of the Bordetella parapertussis adenylate cyclase-										
RT	hemolysin-encoding gene."										
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.										
RM	EMBL; AJ294835; CAB76450.1; -										
SO	LYase.										
	SEQUENCE 1706 AA; 177038 MW; 71750E2D0BBF3B64 CRC64;										

```

Db 294 DYOHGTEQNNPPEADEKEKIVVATGESQMLTRGOLKEYIGQRGEGYFENRABGV- 352
QY 56 PGCSOHMSYGLRPGCSGSDMSYGLRPGCSSOHMSYGLRPGSGOD-----WYGLRPG 109
Db 353 -AGKSLFDPDGGAPGV-----PGCRSK-----SSPDVLETVPASGR-- 391
QY 110 SSFPTGAKKILITVPONVOYDTEQGNGLDOLVKAABELGIEVQREERNMIATAQTSLT 169
Db 392 ---PSLGAVE-----RQDSGIDSIDGYGSRF-----SLGE 419
QY 170 IQTALIGTERGIVLSAPQIDKLLOKTRAGALSAESIYONANK--AKTVLSGIQSLGS 227
Db 420 VSDM-----AAVEAAELEMTQROVLHAGANQDDAEPGVSAGASAHWGQALOGAQAFAA 472
QY 228 VLAGMDLDELALNNSNOHALA-----KAGIELTNSLIENIANSVKTLDFEDEDIOSFGS 291
Db 473 -----ORLVHAIAMLTQFGRAGSTNTPQBAASISAAPFELIGERASSAVALTVS 519
QY 282 KLQNKIGLGTLDKLNKIGCDRAKGLGIDYISGLTSACATLVALYADBNASTAKVAGCFE 341
Db 520 -----GFRGSSRMWAGGFGVAGGAMALOGGIIAAVAGMSLTD--DAPAGOKAAVGA 570
QY 342 LANOVVGNITVAWSYIL-----AQRVAAGLSSTGPVAAALIASTVSLAISPLAFGIADKF 397
Db 571 IALQITGTVEIASISIALALAAARGVTSGLQVAGASGAANAAGALAAALSPMEITGLVOQS 630
QY 398 NHAKSLESYAERFEKLGIDGDNLLAEYQRTGTTIDASVTALINTPLAALIAAGVSAAAA 454
Db 631 HVAQDLKLIAOESSAYEGGALLAOLYRDKTAAGVAVAGSVALSVGAIVSTAAA 667

```

Search completed: March 2, 2001, 11:01:01  
Job time: 471 sec

```
Query Match      5.7%; Score 205.5; DB 2; Length 1706;
Best Local Similarity 22.4%; Pred. No. 0.02;
Matches 107; Conservative 66; Mismatches 193; Indels 111; Gaps 16;

OY   ^C.DRSQHWSTGLRPGSGSDMSYGLLRPGGSSQHWSYC.LRRPGSSQ-----DMSYGLIR 55
      |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:55:42 ; Search time 56.87 Seconds  
(without alignments)  
829.804 Million cell updates/sec

Title: US-09-306-689-13

Perfect score: 3591

Sequence: 1 MATVIDRSGHMSYGLRPGSG.....GLRPGSGSDMSYGLRPGSG 695

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR-66:\*

1: p1r1:\*\*  
2: p1r2:\*\*  
3: p1r3:\*\*  
4: p1r4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2108.5	58.7	953	B30169	leukotoxin A - Pas
2	1748	48.7	935	A35254	leukotoxin A - Pas
3	1264.5	35.2	956	B33389	toxin II - Actinob
4	1248.5	34.8	956	A43834	toxin II - Actinob
5	955.5	26.6	1052	B49219	toxin III - Actino
6	938.5	26.1	1049	SS1784	toxin III - Actino
7	848.5	23.6	1022	I39643	RTX-toxin I - Acti
8	847	23.6	998	T00227	hemolysin A toxin
9	844	23.0	998	I41078	hemolysin A - Escher
10	827	23.0	1055	A37205	leukotoxin A - Act
11	823.5	22.9	1023	LEPCA	hemolysin A - Esch
12	819	22.8	1024	S10056	hemolysin A - Pas
13	554.5	15.4	208	S34238	leukotoxin A - Pas
14	212	5.9	1706	OYBRC	cycloolysin A - Borde
15	206	5.7	1705	S51672	adenylate cyclase
16	178	5.0	2491	A57036	talin - slime mold
17	172	4.8	643	KRHU2	keratin 1, type II
18	166	4.6	638	I53169	cytokeratin 2 - hu
19	166	4.6	645	A44861	keratin, 67k type
20	166	4.6	756	E75590	methyl-accepting c
21	164	4.6	565	I41061	flagellin - Escher
22	163	4.5	622	I37984	keratin 9, type I,
23	161	4.5	990	T38274	probable transcrip
24	157.5	4.4	507	S33192	phase-1 flagellin
25	157.5	4.4	508	A53465	phase 1 flagellin
26	157.5	4.4	2154	F83068	hypothetical prote
27	156	4.3	1190	A82615	surface protein XF
28	155.5	4.3	1045	A39199	endoglucanase B (E
29	155	4.3	936	I40711	sepb protein - Cam

30	153	4.3	2032	2	I39917	hypothetical prote
31	153	4.3	4199	2	S76412	hypothetical prote
32	152	4.2	1467	2	A75564	conserved hypochet
33	151	4.2	641	2	C82206	methyl-accepting c
34	150.5	4.2	893	2	A37284	surface-array prot
35	150	4.2	1822	2	S33441	EF protein - Strep
36	148.5	4.1	2059	2	D82671	surface protein XF
37	147.5	4.1	5627	2	C83339	hypothetical prote
38	147	4.1	1082	2	H81020	serotype-1-specifi
39	146.5	4.1	559	2	F81809	hypothetical DNA-b
40	145	4.0	953	2	S75285	methyl-accepting c
41	144.5	4.0	1113	2	S28925	nuclear pore compl
42	144.5	4.0	3535	2	E83641	probable hemagglut
43	144	4.0	593	1	KRHU0	keratin 10, type I
44	143.5	4.0	424	2	T08093	oleosin homolog ST
45	143.5	4.0	575	2	I40615	flagellin protein

## ALIGNMENTS

```
RESULT 1
B30169
leukotoxin A - Pasteurella haemolytica (serotype 1)
N:Alternate names: lKta protein
C:Species: Pasteurella haemolytica
C:Date: 12-Oct-1989 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C:Accession: B30169; #C32051; S29516
R:Highlander, S.K.; Chidambaram, M.; Engler, M.J.; Weinstock, G.M.
DNA B, 15-28, 1989
A:Title: DNA sequence of the Pasteurella haemolytica leukotoxin gene cluster.
A:Reference number: A30169; MUID:89210283
A:Accession: B30169
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-953 <HIG>
R:Strathdee, C.A.; Lo, R.Y.C.
J. Bacteriol. 171, 916-928, 1989
A:Title: Cloning, nucleotide sequence, and characterization of genes encoding the sec
A:Reference number: A32051; MUID:89123172
A:Accession: C32051
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 947-953 <STR>
R:Lo, R.Y.C.; Strathdee, C.A.; Shewen, P.E.
Infect. Immun. 55, 1987-1996, 1987
A:Title: Nucleotide sequence of the leukotoxin genes of Pasteurella haemolytica A1.
A:Reference number: S29515; MUID:87306837
A:Accession: S29516
A:Molecule type: DNA
A:Residues: 1-741, 'D', 743-953 <LOR>
A:Cross-references: EMBL:M20730; NID:q150492; PIDN:AAA25529.1; PID:q150494
C:Comment: This organism causes bovine pneumonic pasteurellosis (shipping fever).
C:Genetics:
A:Gene: lKta
C:Function:
A:Description: lyses leukocytes
C:Superfamily: hemolysin A; hemolysin A homolog
C:Keywords: calcium binding; cytotoxins; exotoxin; hemolysis; lipoprotein; tandem repe
E:238-784/domain: hemolysin A homolog <HLA>
E:716-807/region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYF]-X)
F:534/Binding site: palmitate (lys) (covalent) #status predicted

Query Match 58.7% ; Score 2108.5 ; DB 1 ; Length 953 ;
Best Local Similarity 51.8% ; Pred. No. 1.3e+94 ;
Matches 484 ; Conservative 1 ; Mismatches 12 ; Indels 437 ; Gaps 2 ;

OY 96 SGSDMSYGLRPGSGSPFKAKIILYIPONYDYTBGNGLDLVKAAELGIEVORE 155
DB 21 SGLHKAGSLTQAGSSL-KTGAKKIILYIPONYGVDFGNGLDLVKAAELGIEVORE 79
OY 156 ERNNTATQTSIGTIQTATIGLTERGIVASAPDIDKLKLTAKAGALGSAESIYONAKAK 215
```

Db	80	ERNIATATQTSLGITQTAIGLTERGIVLSAPQIDKLLQKTKAGALGASASYQNKKK	139
QY	216	TVLSGIOSILGSLVLAGMDLDEALONNSNOHALAKAGLETNLSLENTANSVKTLDGERO	275
Db	140	TVLSGIOSILGSLVLAGMDLDEALONNSNOHALAKAGLETNLSLENTANSVKTLDSEGRQ	199
QY	276	ISOFGSKIONTKIGLGTGDKTKNTGGIDKRGKIGIDVYSGILSGGTAALVADKNASTAK	335
Db	200	ISOFGSKIONTKIGLGTGDKTKNTGGIDKRGKIGIDVYSGILSGGTAALVADKNASTAK	259
QY	336	VGAGFELANQVWGNITTKAVSSYIIAQHVAAGLSTSGVAALIASTVSLAISPLAFAGIAD	395
Db	260	VGAGFELANQVWGNITTKAVSSYIIAQHVAAGLSTSGVAALIASTVSLAISPLAFAGIAD	319
QY	396	KFNHAKSLESYAEBRFFKKLGYDGNMLAEYORGTGTIDASYAINTALAAIAGVSAAAA	454
Db	320	KFNHAKSLESYAEBRFFKKLGYDGNMLAEYORGTGTIDASYAINTALAAIAGVSAAAA	379
QY	455	-----	454
Db	380	SVIASPLALVSGITGYISTILLOYSKOAMEHVAANKIHNKIVEMKNNHCKNYFENGYDA	439
QY	455	-----	454
Db	440	RYLANLQDNMFFLNLNKLEDAEENVAITQOQMDNNIAGLISRLGEKLSKATVDAF	499
QY	455	-----	454
Db	500	EEGKHKADKLVLQDLSANGIIDVNSGKAKTOHILFPTPLPTGTEHREHVOTGKYEYIT	559
QY	455	-----	454
Db	560	KLNINRVDSWKITDGAASSTFDLTNNVQRIEILDNAGNVTKEKTKIAKLGEGDNVF	619
QY	455	-----	454
Db	620	VGSGETLEDGEGYDRVHYSKNGALGTIDATKTEBGSSTVNRFEVETGALHEVISTHT	679
QY	455	-----	454
Db	680	ALVGNREKEIEYRHSNNOHHAGYYTKDTLKAVEEITGSHNDIIFKSGKFDAFNGGDYD	739
QY	455	-----	454
Db	740	TIYGNQDNRLFGKGDDIILDGNGDGFIDGKGNDLHGKGDDIIFVHRKGDGNDIITD	799
QY	455	-----DLTFEKVKNLVLITNSKKEKYTIQNMFREADFAKENVYKATKDE	499
Db	800	SDGNKLSFSDSNLKDOLFEEKVKNLVLITNSKKEKYTIQNMFREADFAKENVYKATKDE	859
QY	500	KIEEIIIGNGERITSKOYDILLAKGNCKITODELSKYVDNVELIKHSKNVTNSLKLISS	559
Db	860	KIEEIIIGNGERITSKOYDILLAKGNCKITODELSKYVDNVELIKHSKNVTNSLKLISS	919
QY	560	VSAFTSSNDSRNVLVAPTSMLDOSLSSIQPARGS	593
Db	920	VSAFTSSNDSRNVLVAPTSMLDOSLSSIQPARAA	953

RESULT 2  
A35254  
leukotoxin A - Pasteurella haemolytica (serotype T10)  
N:Alternate names: Ikta protein  
C:Species: Pasteurella haemolytica  
C:Date: 10-Aug-1990 #sequence revision 15-Nov-1996 #text\_change 18-Jun-1999  
C:Accession: S37145, A35254, S34237, S34235  
R:Lalinson, A.F.; Aitchison, K.; Donachie, W.  
submitted to the EMBL Data Library, September 1993  
A\*:Description: DNA sequence of the leukotoxin A gene from P. haemolytica T10 serotype  
A:Reference number: S37145  
A:Accession: S37145  
A:Molecule type: DNA

A:Residues: 1555 <LA1>  
A:Cross-references: EMBL:Z26247; NID:g400424; PIDN:CA8I206.1; PTD:g400425  
R.Highlander, S.K.; Engler, M.J.; Weinstein, G.M.  
J. Bacteriol. 172, 2343-2350, 1990

J>Title: Secretion and expression of the Pasteurella haemolytica leukotoxin.

A:Reference number: A35254; MUID:90236888

A:Accession: A35254

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 950-955 <HG>

A:Cross-references: GB:M24197; GB:M34943; GB:M34944

R.Lainson, A.F.; Atchison, K.D.; Donachie, W.  
submitted to the EMBL Data Library, June 1993

A>Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T<sub>3</sub>s

A:Reference number: S34235

A:Accession: S34237

A:Molecule type: DNA

A:Residues: 745-955 <LA2>

A:Cross-references: EMBL:Z22884; NID:g311828; PIDN:CA8B0498.1; PTD:g311829

A:Experimental source: serotype T3

A:Accession: S34235

A:Molecule type: DNA

A:Residues: 723-955 <LA3>

A:Cross-references: EMBL:Z22887; NID:g311824; PIDN:CA8B0501.1; PTD:g311825

A:Experimental source: serotype T10

C:Function:

A>Description: attacks cell membranes and causes cell lysis

C:Superfamily: hemolysin A; hemolysin A homologous toxin

G:Keywords: calcium binding; cytotoxicity; hemolysins; lipoprotein; tandem repeats

F:240-786/domain: hemolysin A homology <LVYX>

F:718-809/region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVITYF]-X)

F:718-726/region: repeat

F:727-735/region: repeat

F:736-744/region: repeat

F:745-753/region: repeat

F:754-762/region: repeat

F:763-771/region: repeat

F:772-780/region: repeat

F:781-789/region: repeat

F:792-800/region: repeat

F:801-809/region: repeat

F:556/Binding site: palmitate (lys) (covalent) #status predicted

```

384 VASPIALLVSGITGVISTIIQYSQAMFEHVANKIHNKIYEMEKNNKGKTFENGIDARY 443
QY 455 ----- 454
Db 444 LANLDNMKFLNLNKLQAEERVAITQOOWDSNIGDLAGISRLGKVLGKAYDAFEE 503
QY 455 ----- 454
Db 504 GQHLKADKLVDLSAKGIIDVTNTGEAKTOHILFRPLLPFGTEKREBRVOTGKEYITKL 563
QY 455 ----- 454
Db 564 HINRVDSMOIKDGASSTFDLTNVVQRIQVLELHAENVIKTKETKIYATLGDGDNVFVG 623
QY 455 ----- 454
Db 624 SGTEIDGEGYDRVHYSRNGYALITDATEREGSYTVNRFVESGKALHEGTSTHAL 683
QY 455 ----- 454
Db 684 VGNREEKIEYRHSNNQHAQYTTKDLKAVEEIIQTSNIDIFKSGKPNDAFNGGDVDTI 743
QY 455 ----- 454
Db 744 DGNDRNRLFGGKGDIIIDGNGDDFIDGKGNDLHGKGDILFVHRQGDNDSTESE 803
QY 455 ----- 454
Db 804 GNDKLPSSDNLKDLFEKYNHHLVINTKQEKVTIOMNFEAFKATIQNVYATRDCKI 863
QY 502 EEITIGGERITSKOVDDILAKNGKITDDELKSVNVNVELKSKNVNTSLDKLSSAS 561
Db 864 EEIIGQGERITSKOVDELLEKNGKLAQSELETKRVNDYQNLKYSRDSNSLKLSSAS 923
QY 562 AFTSSNDRNVLAPTSMLDOSLSSLOFARGS 593
Db 924 AFTSSNDRNVLAPTSMLDPSLSSLOFARAA 955

RESULT 3
B33389
N:Altenate names: cytolsin II, RTX-toxin II (Apr11)
C:Species: Actinobacillus pleuropneumoniae
C:Date: 09-Mar-1990 #sequence, revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: B33389; S18853; B43599
R:Chang, Y.F.; Young, R.; Struck, D.K.
DNA 8, 635-647, 1989
A:Title: Cloning and characterization of a hemolysin gene from Actinobacillus (Haemophilus)
A:Reference number: A33389; M0ID:90126233
A:Accession: B33389
A:Molecule type: DNA
A:Residues: 1-956 <CHA>
A:Cross-references: GB:M30602; NID:g141823; PIDN:AAA87232.1; PID:g141825
A:Experimental source: serotype 5
R:Smiths, M.A.; Briatore, J.; Jensen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
Submitted to the EMBL Data Library, July 1991
A:Description: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A:Reference number: S18852
A:Accession: S18853
A:Molecule type: DNA
A:Residues: 1-956 <SMIT>
A:Cross-references: EMBL:X61111; NID:g38939; PIDN:CAA43423.1; PID:g38941
R:Smiths, M.A.; Briatore, J.; Jensen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
Infect. Immun. 59, 4497-4504, 1991
A:Title: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A:Reference number: A43599; M0ID:92040145
A:Accession: B43599
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-27; 948-956 <SM2>
A:Cross-references: GB:X61111; NID:g38939

```

```

C:Comment: This organism causes porcine pleuropneumonia.
C:Genetics:
A:Gene: aprIIA; appA; clyIIA
C:Function:
A:Description: attacks blood cell membranes and causes cell lysis
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytolsins; exotoxin; hemolysis; lipoprotein; tandem repeat
F:243-787/Domain: hemolysin A homology <HVA>
F:719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYIF]-X)
F:557/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match          35.2%  Score 1264.5; DB 1; Length 956;
Best Local Similarity 34.5%  Pred. No. 7.8e-54; Indels 443; Gaps 8;
Matches 320; Conservative 65; Mismatches 100;

QY 104 GLRPGSSFPKTKAKIILYIPQYQYDTFEGNGLODLVKAABELGLEVOREERNIATA 163
Db 34 GLRPGHSL-QNGAKKILIVIPG-YDSGOGNVODLVKAANDGLIEVREERSNLDIA 90
QY 164 QTSIGTQTAIGTERGIVLSAPOIDKLOKT-KAGALGSAESIVONANKAKTVLSGIQ 222
Db 91 KTSFDTQKILGFTDRGIVLFAPODNLKKNPKIGVTLGASSISONIGKANTVLGIG 150
QY 223 SILGSVLAGMDLDELON-NSNOHALAKGIELTNSLIENTANSVKTLDERGEOISQFS 281
Db 151 SILGSVLGVNLNELLONKPNQLELAKAGLELTNELLVGNASSVQTVDAFEOISKLS 210
QY 282 KLOINKIGLGLDKLNIGGLDKAGLGLDVISGLSGATAALVADNASTAKKVGAFGE 341
Db 211 HLOVNVKLGGLSNKLOVLPDLGRKASLGDIISGLSSAGLILADEASTEKKAAGVE 270
QY 342 LANOVNITKAVSSYLLAORVAAGLSSGPAALIASVSLAISPLAFAGIADKFNHAK 401
Db 271 FANQIIGNVTKAVSSYLLAORVAGSLSTGPAALIASVSLAISPLAVADKFKQAD 330
QY 402 SLESYAREPKLGYDGNLAEYORGTGTIDASTAINTALAIAGVSAAA----- 454
Db 331 LKISYSEFQKLGVDGRLLADFHRETGTIDASVTINTALAIAGVSAAGSLVGAP 390
QY 455 ----- 454
Db 391 VALLVAVTGLITILEYSQAMFEHVANKVHRIYEMEKNNKTFEGGDSHLLADIQ 450
QY 455 ----- 454
Db 451 DNMFILNLNKLQAEERVAITQOOWDSNIGDLAISRRTDKISSKAYVDAFEGOHOS 510
QY 455 ----- 454
Db 511 YDSSVQLDNKNKIIINSNTNKTQSVLEFRPLLPGEENRERIOGRKNSYITKLHQRVD 570
QY 455 ----- 454
Db 571 SWVTYDGDASSVDFTVNVQRIAKVFPDAGNIIISKDTKIIANAGANDNVFVSSSTVI 630
QY 455 ----- 454
Db 631 DGGGDHVRVHYSRNGYALVDATEETKESYSVKRYVDSKALHETIATHTQVNGRRE 690
QY 455 ----- 454
Db 691 KIEYREDDREHNTGYTVTSLKSVEEIIIGSFNDIFKSGOPDDVFGNGNVDTIDGNDGD 750
QY 455 ----- 454
Db 751 DHLFGAGDDVDYIDGNGNNPLVGTGNDIISGKNDIIVHKITGDGNDSTIDSGGOKLA 810
QY 455 ----- 454
Db 811 FSDVNLKDLDFEKVDSLEIINQKGEKVRIGNMFLEDDLSTAVNAYKATDRKITEEIIIG 870
QY 508 NGERITSKOVDDILAKNGKITDDELKSVNVNVELKSKNVNTSLDKLSSVSAFPTSSN 567

```



```

0Y 109 GSSPFTAKKILIIYPONVOYDEOENGLODDYKAAEELIEVQREERNINNTAOTSIG 168
Db 48 GRAYOKYQ-KMLVLYIPK--EYDSVSGNGFDFYLKAAEELIQVKYNNRELEYAKHSIG 104
0Y 169 TITAIIGLTERGIYLSAPQIDKLQK-TRAGALGSAE-SIVONANKAKTVLSIGSILG 226
Db 105 TADDFLGLTERGLTFLFAPOLQDFLOKHSKISNNVSSSTGAVSKLASFOTIISIGOSVLG 164
0Y 227 SYLGMDELDELONNSNOHALAKKLETTNLSLEININANKYTLDEPEOISOPGSKONI 286
Db 165 TYLACININENKIIISGGSLELEAGVSLASLAVSINAKGTTTIDAFYTOIONGKULVENA 224
0Y 287 KGLGTLGDKLKNIGG--LDKAGLGLDYISGLSGATAVLYLADKNASTARKVAGAFELAN 344
Db 225 KGLGVGQOLONISGSALEKTLGLDITISLSGVTASFALANKNASTKRVAAGFELSN 284
0Y 345 OYVNGITRAVSSYLLAOPVANGASSTGPVALLINSTVSLIISPLAFAGIDKTRHAKSLE 404
Db 285 OYIGGITRAVSSYLLAOPVANGASSTGPVALLINSSISLSPALAFIRVADNFRNSKEIG 344
0Y 405 SYAEFRKRLAGYDGNLLAEYORGGTIDASVATYNTALAAIAGVSAAAD----- 455
Db 345 EFAEFRKRLAGYDGNLLAEYORGGTIDASVATYNTALAAIAGVSAAAD----- 455
0Y 456 -----LTFEKK-----YKHNLYITNSKKEKVYTIONNFR----- 483
Db 405 LVTGITGLISGLETSKOPMLDHAASKIGNKIDEMWERKYGNGYFENGYDARHRAFLSDSF 464
0Y 484 ---ADFAKEVNPYKATK-----DEKITEELIGONGERITSKOVDDILAKG-----NG 526
Db 465 SLLSEFNQYETERAVLITQORMBEYIGELAG-----ITGK--GDKLSSKAVAVDFQEG 517
0Y 527 KITO---DELSKVVDNVEYELLHSHKNVNS-----LDKLSSVSAFATTSNDR 570
Db 518 KLLKELPPDFSKVY--EDPTGEIDISNOSTJLLKLVPTPL--LTPPTESR 564

```

RESULT 6  
S51784

NALternate names: RTX-toxin IIIA (ApXIIIA)  
N:Species: Actinobacillus pleuropneumoniae  
C:Date: 14-Jul-1995 #sequence\_revision 15-Nov-1996 #text\_change 18-Jun-1999  
C:Accession: S51784  
R:Chang, Y.F.; Shi, J.; Ma, D.P.; Shin, S.J.; Lein, D.H.  
DNA Cell Biol. 12, 351-362, 1993  
A:Title: Molecular analysis of the Actinobacillus pleuropneumoniae RTX toxin-III gene c  
A:Reference number: S51783; MUID:95263992  
A:Accession: S51784  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1049 <CH>  
A:Cross-references: EMBL:L12145; NID:G349605; PIDD:AAA21924.1; PID:q470685  
C:Comment: This organism causes porcine pleuropneumonia.  
C:Genetics:  
A:Gene: apxIIIA  
C:Function:  
A:Description: lyses lung macrophages  
C:Superfamily: hemolysin A; hemolysin A homology  
C:Keywords: calcium binding; cytotoxic; exotoxin; lipoprotein; tandem repeat; thiolester  
F:254-803/Domain: hemolysin A homology <HLYA>  
F:735-861/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYF]-X)  
F:571,702/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 26.1%; Score 938.5; DB 1; Length 1049;  
Best Local Similarity 43.5%; Pred. No. 5e-38;  
Matches 231; Conservative 80; Mismatches 137; Indels 83; Gaps 16;

QY 109 GSFFKTKAKRIILYLIPONTVDTEOGNGLDIVAAAEELGLEVOREERNINATQTSLG 168  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |  
DB 48 GAAYDKRYG-NKTVIPIK-R-EYDSVGSGFGFDVVAEAEEELGLQAVRVNNRLEVAHKSLG 104

```

OY 169 TQIMIGLTERIVISAPOIDKLOK -TKAGALGSAE-SIVONANKKATVUSIOSTIG 226
Db 105 TADQELGTERELTLPAPOLDPFLQKHSKISVNWSSSTGDAVSKLANSQTTISIGISVLG 164:
OY 227 SVIAGMDDEALQONNSOHALAKAGLELTNSLIENTIANSVXTLDEFFEOISQFSKLONI 286
Db 165 TVIAGININEALITISGSELELAEGAVSLASELVSINIAAGTTTIDAFTTQIONFGLAENA 224
OY 287 KULGTLGDKLKING--LDKAGLGIDVISGLISGTAALVILADKNASTAKKVGAFELAN 344
Db 225 KULGEGVROLQINSASALSTKGLGDIISLISGVTREFPALRNKNAASTVVAAFELISN 284
OY 345 QVWGNITKAVSSYIIAORVAATGSLSTGPAAALIASTVSLAISPLAFIADKFNKASLE 404
Db 285 QVIGGITKAVSSYIIAORLRAGLSTTGPAALIASISLAIPLAFILVADNFNNSKIEIG 344
OY 405 STAEFRKLCYDGNULAEYORGTGTIDASVTAINTALAIAGVSAAD----- 455
Db 345 EFAEFRKLCYDGNULAEYEHAGTIDASTITTSALAIAGTAASAGALVGAFTL 404
OY 456 -----LTFREK-----VKHNLVYTNSSKREKVTIONMPRE----- 483
Db 405 LYTGITGLISGLIESKOPMLDHVASKIGNKIDEBEKKYGGKNYFENGIDAHKAFLEDSE 464
OY 484 ---ADFAKEVPNYKATK-----DEKIEIIGONGCERITTSKOVDLILAG-----NG 526
Db 465 SILTSFNKOYETERAVLLITQORMDEYIGELAG-----ITGK--GDKLSSGAAVYDFOEG 517
OY 527 KITTO---BELSKIVNIELKSHKNVNTS---LDKLISSASAFSSDMSDR 570
Db 518 KILEKKPDEFSKVY--FDPTKEIDINSQSTLKEFTVPL--LTPGIESR 564

```

RESULT 7  
 139643  
 RTX-toxin I - Actinobacillus pleuropneumoniae  
 N:Alternate names: hemolysin Apxi  
 C:Species: Actinobacillus pleuropneumoniae  
 C:Date: 19-Jul-1996 #sequence revision 08-Nov-1996 #text\_change 18-Jun-1999  
 C:Accession: 139643; S18769; 139644; S60732; S35781  
 R:Jansen, R.; Briatore, J.; Kamp, E.M.; Gleikens, A.L.; Smits, M.A.  
 Infect. Immun. 61, 3688-3695, 1993  
 A:Title: Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin I (Apxi)  
 A:Reference number: 139641; MUID:93366425  
 A:Accession: 139643  
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB  
 A:Molecule type: DNA  
 A:Residues: 1-1022 <RES>  
 A:Cross-references: EMBL:X73117; NID:G312897; PIDN:CA51548.1; PID:G312899  
 R:Freij, R.; Meier, R.; Gygli, D.; Nicolet, J.  
 Infect. Immun. 59, 3026-3032, 1991  
 A:Title: Nucleotide sequence of the hemolysin I gene from Actinobacillus pleuropneumoniae  
 A:Reference number: S18769; MUID:91348845  
 A:Accession: S18769  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-209, 'AMPYLTIA', 218-373, 'R', 375-561, 'Q', 563-686, 'TC', 688-1022 <FRE>  
 A:Cross-references: EMBL:X52899; NID:G38949; PIDN:CA43081.1; PID:G38950  
 R:Freij, J.; Haldmann, A.; Nicolet, J.; Boffinl, A.; Prentki, P.  
 Gene 142, 97-102, 1994  
 A:Title: Sequence analysis and transcription of the apxi operon (hemolysin I) from Actinobacillus pleuropneumoniae  
 A:Reference number: 139644; MUID:94237497  
 A:Accession: 139645  
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB  
 A:Molecule type: DNA  
 A:Residues: 1-209, 'AMPYLTIA', 218-373, 'R', 375-561, 'Q', 563-686, 'TC', 688-1022 <RE2>  
 A:Cross-references: EMBL:X68595; NID:G505568; PIDN:CA48586.1; PID:G505570  
 R:Tatcon, R.I.; Vazquez-Boland, J.A.; Gutierrez-Martin, C.B.; Rodriguez-Barbosa, I.; Mol, Microbiol. 14, 207-216, 1994  
 A:Title: The RTX haemolysins Apxi and ApxII are major virulence factors of the swine













GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:01:21 ; Search time 32.03 Seconds  
(without alignments)  
700.730 Million cell updates/sec

Title: US-09-306-689-13

Perfect score: 3591  
Sequence: 1 MATVIDRSGHMSYGLRPGSG.....GLRPGSGSDMSYGLRPGGS 695

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2108.5	58.7	953	LKAL_PASHA	P16535 pasteurella
2	1893.5	52.7	953	LKAB_PASHA	P55118 pasteurella
3	1789	49.8	953	LKA3_PASHA	P55116 pasteurella
4	1748	48.7	955	LKAA_PASHA	P55117 pasteurella
5	1585.5	44.2	947	LKTA_PASSP	P55123 pasteurella
6	1265.5	35.2	956	HLXA_ACTSU	Q00951 actinobacil
7	1264.5	35.2	956	RT2A_ACTPL	P15377 actinobacil
8	955.5	26.6	1052	RT32_ACTPL	P55131 actinobacil
9	938.5	26.1	1049	RT31_ACTPL	P55130 actinobacil
10	828.5	23.1	1023	RT11_ACTPL	P55128 actinobacil
11	828.5	23.1	1023	RT12_ACTPL	P55129 actinobacil
12	823.5	22.9	1023	HLX1_ECOLI	P09983 escherichia
13	819.5	22.8	1024	HLXA_ECOLI	P08715 escherichia
14	819	22.8	1050	LKTA_ACTAC	P16462 actinobacil
15	212	5.9	1705	CYAA_BORRE	P15318 bordetella
16	206	5.7	1705	CYAA_BORRE	Q57506 bordetella
17	178	5.0	2491	TALA_DICDI	P54633 dictyostell
18	166.5	4.6	643	K2C1_HUMAN	P04264 homo sapien
19	166	4.6	638	K2Z1_HUMAN	Q01546 homo sapien
20	166	4.6	645	K2Z2_HUMAN	P35908 homo sapien
21	163	4.5	622	K1C1_HUMAN	P35527 homo sapien
22	157.5	4.4	507	FLIC_SALON	Q06974 salmonella
23	155.5	4.3	1045	GONB_CELFI	P26225 cellulomona
24	154	4.3	407	SM41_HEMPU	Q26264 hemlicentrot
25	145.5	4.1	574	FLA3_CAMJE	Q46113 campylobact
26	144.5	4.0	1113	N116_YEAST	Q02630 saccharomyc
27	144	4.0	593	K1C1_HUMAN	P13645 homo sapien
28	143.5	4.0	2869	RBP1_PLAYB	Q00798 plasmodium
29	142	4.0	500	FLJB_SALAE	P5615 salmonella
30	140.5	3.9	414	SVS2_RAT	P32006 rattus norv
31	140.5	3.9	933	SLAP_CAMPE	P33827 campylobact
32	140.5	3.9	2541	TALI_HUMAN	Q94490 homo sapien
33	140	3.9	682	PLIJ_PSEAE	P42257 pseudomonas

## ALIGNMENTS

RESULT	ID	Sequence	Standard	PRT	953 AA.
1	LKAL_PASHA	01-AUG-1990 (Rel. 15, Created)			
AC	P16535	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	LEUKOTOXIN FROM SEROTYPE A1.				
GN	LKTA.				
OS	Pasteurella haemolytica.				
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;				
CC	Pasteurella.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SEROTYPE A1;				
RX	MEDLINE=87306837; PubMed=3040588;				
RA	Lo R.Y.C., Strathdee C.A., Shewen P.E.;				
RT	"Nucleotide sequence of the leukotoxin genes of Pasteurella				
RL	haemolytica A1".				
RN	Infect. Immun. 55:1987-1996(1987).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SEROTYPE A1 / PHL101;				
RX	MEDLINE=89210283; PubMed=2707120;				
RA	Highlander S.K., Chidambaram M., Engler M.J., Weinstein G.M.;				
RT	"DNA sequence of the Pasteurella haemolytica leukotoxin gene				
RL	cluster".				
RN	DNA 8:15-28(1989).				
RN	[3]				
RP	SEQUENCE OF 884-953 FROM N.A.				
RC	STRAIN-SEROTYPE A1 / PHL101;				
RX	MEDLINE=90236888; PubMed=2185213;				
RA	Highlander S.K., Engler M.J., Weinstein G.M.;				
RT	"Secretion and expression of the Pasteurella haemolytica Leukotoxin."				
RL	J. Bacteriol. 172:2343-2350(1990).				
CC	-1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD				
CC	CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY				
CC	DERIVED.				
CC	-1- SUBCELLULAR LOCATION: SECRETED.				
CC	-1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING				
CC	CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC				
CC	ACTIVITY.				
CC	-1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE				
CC	INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).				
CC	-1- PTR: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN				
CC	MODIFIED (BY SIMILARITY).				
CC	-1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				



FT REPEAT 734 739 1.  
 FT REPEAT 743 748 2.  
 FT REPEAT 752 757 3.  
 FT REPEAT 761 766 4.  
 FT REPEAT 770 775 5.  
 FT REPEAT 779 784 6.  
 SQ SEQUENCE 953 AA; 102206 MW; 927FF56CFC884F12 CRC64;

Query Match 52.7%; Score 1893.5; DB 1; Length 953;  
 Best Local Similarity 46.6%; Pred. No. 6,1e-87;  
 Matches 435; Conservative 24; Mismatches 38; Indels 437; Gaps 2;

OY 96 SCSQMSVYCPKRGSSFPKGAIIILYIPONYDTDEGNGLODYLKAAEELGVORE 155  
 DB 21 SCLNRTGSLAAGOSL-KTGAKKIILYIPKDYDTERKNGLODYLKAAEELGVORE 79  
 OY 156 EERNNTATQSTLTOTATGTERGIVLSAPQIDKLLQKTKAGALGSAESIYONANKAR 215  
 DB 80 EGNDAIAKQTSIGTIONVNLGTERGIVLSAPQIDKLLQKTKAGALGSAENLTKFSNMK 139  
 OY 216 TVLSGIQSLGSLVLAGMDLDEALONNSNOHALAKAGLELTNSLIENIANSVKTIDEFGQ 275  
 DB 140 TVLSGIQSLGSLVLAGMDLDEALONNSNELTAKAGLELTNSLIENIANSVKTIDEFGQ 199  
 OY 276 ISQPSKIONIGLGTGDKLNIGSLGKAGIDYISGLSGATATVLAADKNAKSTAK 335  
 DB 200 INQDSKIONVGLSSGLGKGLSGEDTSGLDVSGLSGATATVLAADKNAKSTAK 259  
 OY 336 VQAGELANOVVGNITKAASYLLAQRVAGLSTGCPVALIASVSLASPLAFAGIAD 395  
 DB 260 VQAGELANOVVGNITKAASYLLAQRVAGLSTGCPVALIASVSLASPLAFAGIAD 319  
 OY 336 KRNHAKSLESTAEERKKLGGDGNLAELAEYQRTGTIDASVTAINTALAAIAGVSAA-- 453  
 DB 320 KRNHAKSLESTAEERKKLGGDGNLAELAEYQRTGTIDASVTAINTALAAIAGVSAA-- 379  
 OY 454 ----- 453  
 DB 380 SVIASPIALLVSGITGVISTIIQYSQAMFEHVANKINNKIYEWKNNHGNKTFENGIDA 439  
 OY 454 ----- 453  
 DB 440 RYLANLQDMKFLNMLNKLQAEVYAIITQOQMDNNIGLAGISRLGEVLSKAYVDAA 499  
 OY 454 ----- 453  
 DB 500 EEGHKLKADKLVOLDANGIIDVNSNGAKATODILFRPLTPGTDRERVOGKEYEIT 559  
 OY 454 ----- 453  
 DB 560 KLINRVDWSKITDGAASSTPDLJNVVORIGIELDNAGNVTKETKIVAKLGAGDDNVF 619  
 OY 454 ----- 453  
 DB 620 VSGTTEIDGSGYDRVHNSRGNYGALTIDATKETEQSSYTVNRPVETKALHEGSTHT 679  
 OY 454 ----- 453  
 DB 680 ALVGNREKIEYRHSNNQHHAGYUTKDLKAVEEIIIGTSHNDIFKSGKFNDAFNGGDVD 739  
 OY 454 ----- 453  
 DB 740 TIDGKDGNDRLFGKGDDIIDGNGDDIIDGSKGNDLHGKGDDIIVHROGDGNDIITD 799  
 OY 454 ----- 453  
 DB 800 SDGNKLSFSDSNLKDLPFEKVKHNLVITNSRKEKVTIOWMREAFKAEVANKYATKBE 859  
 OY 500 KIEEIIIGNGERITSKQVDDLIAGKNGKITODELSKVVDNYELLKHSKNVTNSLKLISS 559  
 DB 860 KIEEIIIGNGERITSKQVDDLIAGKNGKITODELSKVVDNYELLKHSKNVTNSLKLISS 919

OY 560 VSAFTSSNDSRNVLVAPTSMLODSSLOFARGS 593  
 DB 920 ASAFITSSNDSRNVLVAPTSMLODSSLOFARAA 953

RESULT 3  
 ID LKA3\_PASHA STANDARD; PRT: 953 AA.  
 AC P5116;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE LEUKOTOXIN FROM SEROTYPE T3.  
 GN LKTA.  
 OS Pasteurella hemolytica.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Pasteurella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SEROTYPE T3;  
 RX MEDLINE=94041617; PubMed=8225575;  
 RA Burrows L.L., Lo R.Y.C., Olah-Witfield E.;  
 RT "Molecular analysis of the leukotoxin determinants from Pasteurella  
 RT hemolytica serotypes 1 to 16."  
 RL Infect. Immun. 61:5001-5007(1993).  
 CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD  
 CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY  
 CC DEFINED.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING  
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC  
 CC ACTIVITY.  
 CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE  
 CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).  
 CC -1- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN  
 CC MODIFIED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: U01216; AAB36691.1;  
 DR HSSP: P02392; ICTF.  
 DR INTERPRO: IPR001343;  
 DR PFM: PF00353; hemolysincabind; 1.  
 DR PROSITE: PS00330; HEMOLYSIN-CALCIUM; 2.  
 KW Hemolysis; Toxin; Cytolysis; Cytolextin; Repeat; Calcium;  
 KW Transmembrane; Lipoprotein; Palmitate.  
 FT TRANSMEM 229 249 POTENTIAL.  
 FT TRANSMEM 297 318 POTENTIAL.  
 FT TRANSMEM 381 401 POTENTIAL.  
 FT DOMAIN 734 784 6 x REPEATS, GLY-RICH.  
 FT REPEAT 734 739 1.  
 FT REPEAT 743 748 2.  
 FT REPEAT 752 757 3.  
 FT REPEAT 761 766 4.  
 FT REPEAT 770 775 5.  
 FT REPEAT 779 784 6.  
 SQ SEQUENCE 953 AA; 101948 MW; FDBDCE2FDC85FDE2 CRC64;

Query Match 49.8%; Score 1789; DB 1; Length 953;  
 Best Local Similarity 44.3%; Pred. No. 9,2e-82;  
 Matches 413; Conservative 38; Mismatches 39; Indels 442; Gaps 2;  
 OY 104 GLRPGSSFP-----KTGAKKIILYIPONYDTDEGNGLODYLKAAEELGVORE 157  
 DB 22 GLNRAGSLTQAGQTLKNGVKKIILYIPKDYKTDSSGNGLODYLKAAEELGVORE 81



[illegible]

```

CC Pasteurella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE T10.
RX MEDLINE=96425875; PubMed=8828217;
RA Latson F.A., Murray J., Davies R.C., Donachie W.;
RT "Characterization of epitopes involved in the neutralization of
RL Pasteurella haemolytica serotype A1 leukotoxin."
RM Microbiology 142:2499-2507(1996).
CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC DEFINED.
CC CC -1- SUBCELLULAR LOCATION: SECRETED.
CC CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY.
CC CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
CC INVOLVED IN PORE FORMATION BY THE CTOTOXIN (BY SIMILARITY).
CC CC -1- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN
CC MODIFIED (BY SIMILARITY).
CC CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: Z26247; CAA81206.1; -.
DR INTERPRO: IPR001343; -.
DR PFAM: PF00035; hemolysinCblnd; 1.
DR PRINTS: PR00313; CABNDNGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCIUM; 4.
KW Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
KW Transmembrane; Lipoprotein; Palmitate.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 383 403 POTENTIAL.
FT TRANSMEM 736 786 6 X REPEATS, GLY-RICH.
FT DOMAIN 736 741 1.
FT REPEAT 745 750 2.
FT REPEAT 754 759 3.
FT REPEAT 763 768 4.
FT REPEAT 772 777 5.
FT REPEAT 781 786 6.
SQ SEQUENCE 955 AA; 102187 MW; B60F2DB8168BCAF CRC64;

Query Match 48.7%; Score 1748; DB 1; Length 955;
Best Local Similarity 43.5%; Pred. No. 1e-79;
Matches 405; Conservative 41; Mismatches 44; Indels 442; Gaps 2;

QY 104 GLRGSGSF-----KGAKKIILYPONVYDEQNGSLDIDLVAKAEEGLIEVQREER 157
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24 GLNRAGGSLTPAQGTLLNKGAKKIILIPKDVIRKYDSGGSGNGDLTVKAEDELGIENVQKEG 83

QY 158 NNMTAQTSLGTITQAIGTERGIYLSAPQIDKLQTKAGCALGSASIVONANKATV 217
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 84 NDIAKAQTSLTIONVGLTEFGIYLSAPQDKLQKNKVCAQLSSSESTQNFSQAKTV 143

QY 218 LSGIOSILSVLAGMDIDEALQNNNSHQALAKAGIELNLSILENTANSVKTLDEGEQIS 277
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 144 LSGVGQSNRYVLAGMDIDEALQNESDQLTLKAGIELNLSLEINTANSVOTLDATSDQIS 203

QY 278 QFGSKLQNIKIGTGDKLNKIGGLDKAGLGELDIVSGLSGATAALVADYNASTAKKVG 337
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 204 QFGSKLVNKKLGALGXDKLNKIGGLDKAGLGELDIVKSRLLSGATVALVADKDASTAKKV 263

QY 338 AGFELANOVGNITRKAVSYLLAQRVAAGLSSTGFVAALISTVSLAISPLAFGIADKF 397
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 264 AGFELANOVGNITRKAVSYLLAQRVAAGLSSTGFVAALISTVSAVAIVATISPLFAIGADF 323

```



```

OY 398 NNAKSLAESYAEERFKKLGYYDGDNLAEYORGSTGTIDASVTAINTALAIAGVSAANA----- 454
DB 324 DAKSLKLENTAEERFKKLGEGDLSLAERYONGTGTIDASVTAINTALAIAGVSAANAAGSV 383
OY 455 ----- 454
DB 384 VASPIALVSGITGYISTILQYSKOAMFEHVANKITHNKIVEMKNNCKNFENGVDARY 443
OY 455 ----- 454
DB 444 IANLQDNMFKLLNKLKLOAERVIAITQOQMDSNIGDLAISRLGKVKLSGRAVYDAFEE 503
OY 455 ----- 454
DB 504 GOHLKADKLVOILDSAKGIIDVTNTGBAKTOHLLFPRLTLPSTGERKERVQTKYEYITKL 563
OY 455 ----- 454
DB 564 HINRVDSMOIKDGAASFTFDLTNNVQRIQVELDHAENVIKTKETKIVATLGDGDDNVFVG 623
OY 455 ----- 454
DB 624 SGTTEIDGEGYDRVYHSGNYGALLTIDATKETEGSYTVNRFVESGKALHGGSTHTAL 683
OY 455 ----- 454
DB 684 VGNREKLEIYRHSNNQHNAGYVTKDTLAKAVEIIGTSHNDIFKSGKFNDAFNGGQGVDTI 743
OY 455 ----- 454
DB 744 DQNDGNRLFGCGKDDIIDGNGDDEFIDGCKNDLLHGKGGDITVYHROGDNDSITSE 803
OY 455 ----- 454
DB 804 GNDKLSFSDSNLKDLEFEKVNHLVITNTKQEKVTIQNMFRABEPAKTIQNVATRDXTI 863
OY 502 ERTIONCERINSKOVDDLIAGKNGKITODELSKYVDNVEILKHSKNVTNSIDKLISVS 561
DB 864 ERTIONCERINSKOVDDLIAGKNGKITODELSKYVDNVEILKHSKNVTNSIDKLISVS 561
OY 562 AFTSSNDRSNVLVAPTSMLDQSLSSLOFARCS 593
DB 924 AFTSSNDRSNVLVAPTSMLDQSLSSLOFARCS 955

RESULT 5
ID LKTA_PASSP STANDARD: PRF: 947 AA.
AC P55123.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LEUKOTOXIN.
GN LKTA.
OS Pasteurella haemolytica-like sp. (strain 5943B).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93239320; PubMed=8478098;
RA Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RT "Molecular characterization of a leukotoxin gene from a Pasteurella
RT haemolytica-like organism, encoding a new member of the RTX toxin
RT family."
RL Infect. Immun. 61:2089-2095(1993).
CC -I- FUNCTION: VIRULENCE FACTOR WHICH IS CYTOTOXIC FOR LEUKOCYTES BUT
CC IS NOT HEMOLYTIC.
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY.
CC -I- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE

```

```

CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -I- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN
CC MODIFIED (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: L12148; AAA16444.1; -.
DR INTERPRO: IPR001343; -.
DR PFAM: PF00353; HemolysinCabin; 1.
DR PRINTS: PR00313; CABDNDRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCIUM; 4.
KM Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
KT Transmembrane; Lipoprotein; Palmitate.
FT TRANSMEM 154 170 POTENTIAL.
FT TRANSMEM 312 333 POTENTIAL.
FT TRANSMEM 393 414 POTENTIAL.
FT DOMAIN 625 780 1.
FT REPEAT 625 630 7 X REPEATS, GLY-RICH.
FT REPEAT 730 735 2.
FT REPEAT 739 744 3.
FT REPEAT 748 753 4.
FT REPEAT 757 762 5.
FT REPEAT 766 771 6.
FT REPEAT 775 780 7.
SQ SEQUENCE 947 AA; 101559 MW; 9744F06395EF5BED CRC64;

```

```

Query Match 44.2%; Score 1585.5; DB 1: Length 947;
Best Local Similarity 39.7%; Pred. No. 1.1e-71;
Matches 370; Conservative 56; Mismatches 70; Indels 437; Gaps 3;

OY 96 SSGQMSYGLRPGSSFPRTGAKTILYIPONYQYDEOGNGLQDLVKAEBLGIEVORE 155
DB 17 SGLHKTGSLNAGOSL-KACAKKILYIPKDYEDSGRGLQDLVKAEDLGIEVORE 75

OY 156 EENNATATQTSIGTQTAIGLTERGIVLSAPQIDKLOKTRAGQALGSASIVQANAK 215
DB 76 ERNGIATQNSLSTIQNLIGSERGVLSAPQIDKLOKTRAGQALGSASIVQANAK 135

OY 216 TVLSIGQSLGSLVLAGMDLDEALQNSNOHALAKAGLELNSLENINSYKTDDEFQEO 275
DB 136 TLLSGTQSLGSLVLAGMDLDEALQNSNOHALAKAGLELNSLENINSYKTDDEFQEO 195

OY 276 ISQFGSKLQNIKGLTGLDKLKNIGGLDKAGLGLDIVISGLLSGATPAALVLADKKNASTAKK 335
DB 196 ISQFGSKLQNIKGLTGLDKLKNIGGLDKAGLGLDIVISGLLSGATPAALVLADKKNASTAKK 255

OY 336 VGAGFELANQVGNITTKVAVSYIIAQRVAAGLSSGPVAAALIASVLSAISPFAAGIAD 395
DB 256 VGAGFELANQVGNITTKVAVSYIIAQRVAAGLSSGPVAAALIASVLSAISPFAAGIAD 315

OY 396 KFNNAKSLAESYAEERFKKLGYYDGDNLAEYORGSTGTIDASVTAINTALAIAGVSAANA- 454
DB 316 KFNNAKSLAESYAEERFKKLGEGDLSLAERYONGTGTIDASVTAINTALAIAGVSAANAAG 375

OY 455 ----- 454
DB 376 SLVGAPIALVSGITGYISTILQYSKOAMFEHVANKIHDKIYDMEKHNKKNYFENGYS 435
OY 455 ----- 454
DB 436 RYLADLQDNMROLNKLKLOAERVIAITQOQMDSNIGDLAISRLGKVKLSGRAVYDAF 495
OY 455 ----- 454
DB 496 EBGKIKADTFVQILDSAKGIIIDVTNTGBAKTOHLLFPRLTLPSTGERKERVQTKYEYIT 555

```

QY 455 ----- 454  
 Db 556 KLINIRVDSMKITDGAATNSFDLNNVQRIEILDHADNVTKEKTIANLGDGNDVF 615  
 QY 455 ----- 454  
 Db 616 IGSQTEVDGNGLDVRVHRSRGDYGALITDATNESVGSYTVKRFVETGALHEVTATQS 675  
 QY 455 ----- 454  
 Db 676 VLVGSREKEIEYRHSNNTQAGYTTTDLKASVEEIICTSRNDIFKSGKFDPAFHGGDGV 735  
 QY 455 ----- 454  
 Db 736 NIDGNAGNDRLFGGKGFIDIDGGGDGDFIDGGGDDILHGKGNLICTYKGNDSISDS 795  
 QY 455 ----- 500  
 Db 796 GGNDRLSFADSNLKDLEFEYVNHMLMTNVEKVTIONFREADYAKTYHNTQATADEK 855  
 QY 501 IEEIIGNGERITSKOYDDIANGNGKITODELSKYVDNVELKHSNNTNSIDKLISV 560  
 Db 856 IEEIIGNGERITSKOYDDIANGNGKITODELSKYVDNVELKHSNNTNSIDKLISV 560  
 QY 561 SAFTSSNDSNRNVLAFTSMLODLSLQFARGS 593  
 Db 916 GAFASSNDNRNVGCVPTSLYEHT-QSVQFVRAA 947

## RESULT 6

HLVA\_ACTSU STANDARD; PRT; 956 AA.

ID AC 000951;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE HEMOLYSIN (CYTOLYSIN II) (CLY-IIA) (HLV-IIA) (CYTC) (APPA).

GN APPA OR CLYIIA OR HLVIIA OR CYTC.

OS Actinobacillus suis.

OC Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Actinobacillus.

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-3714;

RX MEDLINE-92267623; PubMed-1587585;

RA Burrows L.V., Lo R.Y.;

RT "Molecular characterization of an RTX toxin determinant from

Actinobacillus suis."

RL Infect. Immun. 60:2166-2173(1992).

CC -1- FUNCTION: ONE OF THE VIRULENCE FACTORS OF A.SUIS. MIGHT BE A

SECRETED CYTOTOXIN, POSSIBLY THE EXTRACELLULAR HEMOLYSIN.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING

CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC

ACTIVITY.

CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE

INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).

CC -1- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN

MODIFIED (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M90440; AAA21918.1;

DR INTERPRO: IPR001343;

DR PFAM: PF00353; hemolysinCbind; 1.

DR PROSITE: PS00330; HEMOLYSIN\_CALCIUM; 1.  
 KM Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;  
 KW Transmembrane; Lipoprotein; Palmitate.  
 FT TRANSMEM 238 254  
 FT TRANSMEM 302 320  
 FT TRANSMEM 383 406  
 FT DOMAIN 719 799  
 FT REPEAT 719 724  
 FT REPEAT 728 733  
 FT REPEAT 737 742  
 FT REPEAT 746 751  
 FT REPEAT 755 760  
 FT REPEAT 764 769  
 FT REPEAT 773 778  
 FT REPEAT 782 787  
 FT REPEAT 794 799  
 SO SEQUENCE 956 AA; 102453 MW; 3415FF1D7AD4365 CRC64;

## Query Match

Best local Similarity 35.2%; Score 1265.5; DB 1; Length 956;

Matches 320; Conservative 65; Mismatches 100; Indels 443; Gaps 8;

Db 104 GLRPGSSFPKGTGAKKIIYIPONYOYDTEOGNGLODLYKAABELGIEVOREERNIATA 163

Db 34 GLNOTGHSI-QNGAKKILIIYIPOG--YDSGONGIIDLVAANDLGIEVREERSNIDIA 90

QY 164 QTSIGTQTAIGTERIVTASPOIDKLOKT-KAGALGSARSYONAKKATVLSG 222

Db 91 KTSFDYTKILGFTDRKIVTAPOLDNLKPKRGITGTLGASASSISONGIKATVLSG 150

QY 223 SILGSLVAGMDLDEALQN-NSNOHALAKAGLENTSLENIANSVYTLDFEGQISQFS 281

Db 151 SILGSLVAGVNLNELLQNKPNOLAKAGLELTNELVGLASSVQTVDAFAQISKLS 210

QY 282 KLDNKGGLTGLKLNKIGGLRAGLGLDVISGLSGATAVLADKNASTARKVAGFE 341

Db 211 HLNVKGLGSLSKRLQNLPLGRASIGLIDTISGLSGASGLILADKKAATEKKAAGVE 270

QY 342 LANOVGNITKRAVSVYLAQRAVAAAGSSGPAALIASVSLAISPLAFAGINDKHNK 401

Db 271 PANQITGNVTKRAVSVYLAQRAVAAAGSSGPAALIASVSLAISPLAFAGINDKHNK 330

QY 402 SLESYAEFFKLGVDGNDLAEYORGTTIDASVTAINFLAIAAGVSAANA----- 454

Db 331 LINSYSERFQKLGTDGRLADPHRETGITDASVTITNFLAIAISGVGASAGSLVAP 390

QY 455 ----- 454

Db 391 VALLVAGVTGLITILEYSKQAFEHYANKHVDRIWEMKHNKYPEQGYDSRHLADQ 450

QY 455 ----- 454

Db 451 DNMFELINKLEQAEVVAITQQRMDNIGDLAISRTQISSKAYDAFEGNISP 510

QY 455 ----- 454

Db 511 SHPYSIDNKNKGIINISNTKRTQSVLFRTPLTPGEBENERIQEKNSTYTKLHQVD 570

QY 455 ----- 454

Db 571 SMTVTVGADASSVDFTNVVQRIAVKFFDAGNIIESKDTKLIANLGGANDVFGSSTVYI 630

QY 455 ----- 454

Db 631 DGGDGHDRVHYSNGEYALVIDATAETEGSSYKRYVGDSCALHETIATHQTVGNREE 690

QY 455 ----- 454

Db 691 KIEYREDDRFHTGTYVDSLKSVEELISQFNDIFKSGQFDVFFHGNGVDTIDGNDG 750

QY 455 ----- 454

[illegible]

Query Match	Best Local Similarity	Score	DB 1	Length
Matches 320; Conservative 65; Mismatches 100; Indels 443; Gaps	34.58%	1264.5	9.4e-56;	956;
Transmembrane; Lipoprotein; Palmitate.				
FT TRANSMEM 233 256			POTENTIAL.	
FT TRANSMEM 266 323			POTENTIAL.	
FT TRANSMEM 361 406			POTENTIAL.	
FT DOMAIN 719 787			8 X REPEATS, GLY-RICH.	
FT REPEAT 719 724			1.	
FT REPEAT 728 742			2.	
FT REPEAT 737 743			3.	
FT REPEAT 746 751			4.	
FT REPEAT 755 760			5.	
FT REPEAT 764 769			6.	
FT REPEAT 773 778			7.	
FT REPEAT 782 787			8.	
SQ SEQUENCE 956 AA; 102531 MW; BDBCABADFL4A641 CRC64;				
Query Match	35.28%	Score 1264.5;	DB 1;	Length 956;
Best Local Similarity	34.58%	Pred. No. 9.4e-56;		
Matches 320; Conservative 65; Mismatches 100; Indels 443; Gaps				
104 GLPGSSSPFKGAKKILITYPONTQYDTDEGNGNGLDYLKAAEELGIEVOREERNIATA 163				
34 GLTQGHSL-QNGAKKLLIPIQG--YDSGGNGYQDLVKRANDIGLEWREESNDIA 90				
164 QTSICTQTAIGLTERGIVLSAPQIDKLQRT-KAGALGSAESIYONAKRATVLSGIQ 222				
91 KTSPTQKILGFTDRGIVLFAPOLDNLKKPKIKGTLGSSASSISQNIKANVTLGIG 150				
223 SLTSGVLAMGMDLELQNSQNALAKAGLELTNSLENIANSVKTLDERGEDISQFS 281				
151 SLTSGVLGVMNLNELLQKRPQOLELAAAGLELTNELVGNLASSVQYVDAEQISKLS 210				
282 KIQNKGLGTLDKKNKNGGLDKAGLDGIVTISGLSGATLVADNASTAKKVGAGFE 344				
211 HIQNWKGLGSLNKNQNPDLCKASLGDIISGLLSGASALILADDEASTEKKAAGVE 270				
342 LANQVGNITRAVSSYLLAQRVAGLSSTGVAALLASTVSLATSPLAAGIADKFNHAK 401				
271 FANQIIGVNTKAVSSYLLAQRVAGLSSTGVAALLASTVSLATSPLAAGIADKFNHAK 330				
402 SLESAERFKIIGYQGNULAEYQGTITDASTAINTALALAGVSAANA----- 454				
331 LKSYSERQKLGYGDRILDHFHRETGTIDASTYTTINTALAAISGVGAASAGSLVGP 390				
455 ----- 454				
391 VALLVAGVTGLITTLLEYSKQAMFEHVANKYHRIVEKKNKRYDEQGYDSRHLADQ 456				
455 ----- 454				
451 DMKKFLINKELOAEVVAITQQRWDQIDLAAISRDTQISSKRAYDAPEEGHQS 510				
455 ----- 454				
511 YDSVOLDNKNKGIINISNTNKTQSVLFRPLTPGEEENREROEGRNSYITKLHIQRYD 570				
455 ----- 454				
571 SMTVTGDGASSVDFTNVVQRIAKFDDAGNIIESKDTKIITANGAGNDNVGSSSTTVI 630				
455 ----- 454				
631 DGGGDHNRVHYSRGEGALVYIDATAEETKGSYSVKRYVGSCKALHETIATHQTVNGREE 690				
455 ----- 454				
691 KLEVRREDDREHTGYTVLDSLKSVEELISQFNDIFKSGQFDVFGHNGVDTIDGNDGD 750				
455 ----- 454				
751 DHLFGAGDDVYIDGGNGNPLVGGGNDIISGKQNDIYVHKGTGDNDSITDSCGQDLA 810				
455 ----- 454				

```

Db      811 FSDVNLKDLTFKRVDSLEIINQGEKVRIGNMFLEDDLSTVANYKATDRKIEETIGK 870
OY      508 NGRIRTSKQVDDLIANKNGKITODELSKVVDNELLKHNNVNSLKLSSVSAPFTSSN 567
Db      871 GGEIRTSQVDDKLKESGNNQISAEALSKVVDNTVTSKRONVNSLKLSSVSGFTSSS 930
OY      568 DSRNVL--VAPTSMDQSLSSLOFARGS 593
Db      931 DFRNNGTGVPPSS-IDVS-NNIGOLARAA 956

RESULT 8
RT32_ACTPL STANDARD; PRT; 1052 AA.
ID RT32_ACTPL STANDARD; PRT; 1052 AA.
AC P55131;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RTX-III TOXIN DETERMINANT A FROM SEROTYPE 8 (APX-IIIA) (CYTOLYSIN DE IIIA) (CLY-IIIA).
GN APXIIIA OR CLYIIIA OR RTX OR PTXA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Actinobacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-405 / SEROTYPE 8;
RX MEDLINE-95012630; PubMed-7927703;
RA Jansen R., Briatore J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J., Smit M.A.;
RT "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX) operon: characterization of the ApXIII operons.";
RL Infect. Immun. 62:4411-4418(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE 8;
RX MEDLINE-93162836; PubMed-8432615;
RA Jansen R., Briatore J., Kamp E.M., Gielkens A.L.J., Smit M.A.;
RT "Cloning and characterization of the Actinobacillus pleuropneumoniae RTX-toxin III (ApXIII) gene.";
RL Infect. Immun. 61:947-954(1993).
CC -1- FUNCTION: DOES NOT HAVE HEMOLYTIC ACTIVITY BUT SHOWS A STRONG CYTOTOXICITY TOWARDS ALVEOLAR MACROPHAGES AND NEUTROPHILS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC ACTIVITY (BY SIMILARITY).
CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -1- PTM: PALMITOYLATED BY APXIIIC. THE TOXIN ONLY BECOMES ACTIVE WHEN MODIFIED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X80055; CAA56358.1;
CC EMBL: X68815; CAA48711.1;
CC INTERPRO: IPR001343;
DR PFAM: PF00353; hemolysin_cabind; 2.
DR PROSITE: PS00330; HEMOLYSIN_CALCIIUM; 3.
KW Toxin; Cytolysis; Cytolextin; Repeat; Calcium; Transmembrane; Lipoprotein; Palmitate.
FT *TRANSMEM 248 265
FT *TRANSMEM 275 334
FT *TRANSMEM 372 418
FT *DOMAIN 754 859
FT 7 X REPEATS, GLY-RICH.

```

```

FT REPEAT 754 759 1.
FT REPEAT 763 768 2.
FT REPEAT 772 777 3.
FT REPEAT 781 786 4.
FT REPEAT 790 795 5.
FT REPEAT 799 804 6.
FT REPEAT 808 813 7.
FT REPEAT 827 832 8.
FT REPEAT 836 841 9.
FT REPEAT 845 850 10.
FT REPEAT 854 859 11.
SQ SEQUENCE 1052 AA; 112809 MW; F83AFE25A6FD8758 CRC64;

Query Match 26.6%; Score 955.5; DB 1; Length 1052;
Best Local Similarity 44.1%; Pred. No. 2,2e-40;
Matches 234; Conservative 61; Mismatches 133; Indels 83; Gaps 16;

OY 109 GSSEPKTGAKKIILYIPONYQYDTEQNGLODLYKAAELGIEVOREERNNTAQTSLG 168
Db 48 GKAVQKYG-NKLVLYIRK--EYDGSVGNFEDLYKAAELGIQVKNRNELEVAHKSILG 104
OY 169 TIQTATLTERGIVLAPQIDKLLOK-TRAGQALGSAE-SIVONANKAKTVLSIGSILG 226
Db 105 TADQFLGLTERGLTLPAPQIDQFLQKHSKISNVYSGTGDVASKLAKSQFTISGIVLG 164
OY 227 SYLAGMDLEALQNNQNALAKAGLELTNSLIENANSVKTDERGEOISQPSKLONI 286
Db 165 TVLAGINLNEALITSGSELELAERAGVSLASELYSNIAKGTITIDAFQTIONFGKLYENA 224
OY 287 KGLATLGDKLKNIIG--LDKAGLGLDVISGLSGATPAALVLAADKNASTAKKVGAGELAN 344
Db 225 KGLGAGVQRLQNTISGALSSTKGLDITSLSGVASFALAKNNKSTKVAAGFELSN 284
OY 345 QVVGNTTKAVSYTLAORVAAGLSSTGPVAALIASVSLAISPARGIDKRNHAKSLE 404
Db 285 QVIGITKAVSYTLAORLAAGLSTGPAPALIASSISLISPALFARVADNENRSKEIG 344
OY 405 SYDERFKLLDYDGNLAEYQRTGTIDASVTAINPALIAGVSAAD----- 455
Db 345 EPRERFKLLDYDGNLAEYQRTGTIDASVTAINPALIAGVSAAD----- 455
OY 445 EPRERFKLLDYDGNLAEYQRTGTIDASVTAINPALIAGVSAAD----- 455
Db 445 EPRERFKLLDYDGNLAEYQRTGTIDASVTAINPALIAGVSAAD----- 455
OY 456 -----LTFEK-----VKHNLVITNSKKEKVTIONMFR----- 483
Db 405 LVYGITGLISGLEFSKOPMLDVASKIGNKIDEMEKKYKNFYENGYDARHKAFLDSP 464
OY 484 ---ADFAKEVPNKATK-----DEKEELIIGONGEKITSKQVDDLIANK-----NG 526
Db 465 SLSSFEKKQYETERAVLITQORNDYIGELAG---ITGK--GDKLSSGKAYVDYFQEG 517
OY 527 KITO---DELSKVVDNDELKHSKNVTNS---LDKLSSVSAPFTSSNDR 570
Db 518 KLEKKRDKDPSKYV--FDPRTKGEIDISNGSTLTKKVTPL--LTGTETSR 564

RESULT 9
RT31_ACTPL STANDARD; PRT; 1049 AA.
ID RT31_ACTPL STANDARD; PRT; 1049 AA.
AC P55130;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RTX-III TOXIN DETERMINANT A FROM SEROTYPE 2 (APX-IIIA) (CYTOLYSIN DE IIIA) (CLY-IIIA).
GN APXIIIA OR CLYIIIA OR RTX OR PTXA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Actinobacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE 2;
RX MEDLINE-93263992; PubMed-8494611;
RA Chang Y.-F., Shi J., Ma D.-P., Shin S.J., Iain D.H.;

```

RT "Molecular analysis of the *Actinobacillus pleuropneumoniae* RTX  
 toxin-III gene cluster.";  
 RL DNA Cell Biol. 12:351-362(1993).  
 RN [2]  
 RP SEQUENCE OF 828-1049 FROM N.A.  
 RC STRAIN=1536 / SEROTYPE 2;  
 RX MEDLINE=95012630; PubMed=7927703;  
 RA Jansen R., Briatore J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,  
 RA Smits M.A.;  
 RT "Genetic map of the *Actinobacillus pleuropneumoniae* RTX-toxin (Apx)  
 operon: characterization of the ApxIII operon.";  
 RL Infect. Immun. 62:4411-4418(1994).  
 CC -1- FUNCTION: DOES NOT HAVE HEMOLYTIC ACTIVITY BUT SHOWS A STRONG  
 CC CYTOTOXICITY TOWARDS ALVEOLAR MACROPHAGES AND NEUTROPHILS.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING  
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC  
 CC ACTIVITY (BY SIMILARITY).  
 CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE  
 CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).  
 CC -1- PTM: PALMITOYLATED BY APTIIC. THE TOXIN ONLY BECOMES ACTIVE WHEN  
 CC MODIFIED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: L12145; AAA32924.1; -;  
 DR EMBL: X80056; CAB37652.1; ALT-SEQ.  
 DR INTERPRO: IPR001343; -;  
 DR PFAM: PF00353; hemolysinCbind; 2.  
 DR PROSITE: PS00330; HEMOLYSIN-CALCIUM; 3.  
 KW Toxin; Cytolysis; Cytotoxin; Repeat; Calcium; Transmembrane;  
 KW Lipoprotein; Palmitate.  
 FT TRANSMEM 154 170 POTENTIAL.  
 FT TRANSMEM 315 331 POTENTIAL.  
 FT TRANSMEM 397 413 POTENTIAL.  
 FT DOMAIN 753 858 11 X REPEATS, GLY-RICH.  
 FT REPEAT 753 758 1.  
 FT REPEAT 762 767 2.  
 FT REPEAT 771 776 3.  
 FT REPEAT 780 785 4.  
 FT REPEAT 789 794 5.  
 FT REPEAT 798 803 6.  
 FT REPEAT 807 812 7.  
 FT REPEAT 826 831 8.  
 FT REPEAT 835 840 9.  
 FT REPEAT 844 849 10.  
 FT REPEAT 853 858 11.  
 SQ SEQUENCE 1049 AA; 112491 MW; F99846BFD4E5CE72 CRC64;

Query Match 26.1%; Score 938.5; DB 1; Length 1049;  
 Best Local Similarity 43.5%; Pred. No. 1.5e-39;  
 Matches 231; Conservative 80; Mismatches 137; Indels 83; Gaps 16;

QY 109 GSPFKTGAKKIIYIPQNYOYDFGNGLODLYKAEELEIGVOREERNINATQTSIG 168  
 DB 48 GKAVQKYG-NKLVLVYIPK--EYDGSVGNFDPDYKAAAEELGIQVYVNNNEIEVAHKSILG 104  
 QY 169 TITQAIIGTERGIYLSAPQIDKLQK-TRAGQALGSAE-SIVONANKAKTYLVSIGIOTILG 226  
 DB 105 TADQGLTGERGLTFLAPQLODFLOKHSKISNVGSGTDGAVSKAKSQTITISIGIOTILG 164  
 QY 227 SYLAGMDLDEALONNSNOHALAKAGLETNSLIENIANSVKTLDEFGQISOFGSKLONT 286  
 DB 165 TYLAGININLEAIISGSELELAAGVSLASLAVSNIAKGTITIDAFQTIONFGKLAENA 224

QY 287 KGLGTGDKLNKNG--LDKAGLGLDVISGLSGATAALVLADNNAKRAKVGAGFELAN 344  
 DB 225 KGLGCVGRQLOINSSALSASTKGLGDISSILSGVTRSFALRNNAKRAKSTVAGFELSN 284  
 QY 345 QVGNITKAVSSYIIAQRVAAAGLSSTGVPVAAIIASTVSLAISPLAFACIADFNHANSLE 404  
 DB 285 QVIGITKAVSSYIIAQRVAAAGLSSTGVPVAAIIASTVSLAISPLAFACIADFNHANSLE 344  
 QY 405 SYAEFRKRLGVDGDLAEYGRGTIDASVTAINFATAIAGVSAAD----- 455  
 DB 345 EFAEFRKRLGVDGDLAEYGRGTIDASVTAINFATAIAGVSAAD----- 455  
 QY 456 -----LPEEK-----YKHNVLITNSKKREKVTIONMFE----- 483  
 DB 405 LVTGTTGLISGLTERSPMDHVAASKIGNKIDEMKRYGKNYPENGIDARHKAFLDSF 464  
 QY 484 ---ADFAKEVPNYATK-----DEKIEIIGONGERITSRQVDDLIAGK-----NG 526  
 DB 465 SLSSFNKQYETERAVLITQGRWDEYIGELAG-----ITGK--GDKLSSGRAVYDFQEG 517  
 QY 527 KITQ---DELSKYVDNYELLHRSKRVTS---LDKLSSVSATFSSDSR 570  
 DB 518 KLEKKRPDFSKV--FDPTRGEIDISNSQSTLKLFTPL--LTPGTESR 564

RESULT 10  
 RT11.ACTPL STANDARD; PRT: 1023 AA.  
 AC P55128;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE RTX-I TOXIN DETERMINANT A FROM SEROTYPES 1/9 (APX-IA) (HEMOLYSIN IA)  
 DE (HLY-IA) (CYTOLYSIN IA) (CLY-IA).  
 GN ApxIA OR CLYIA OR HLYIA.  
 OS *Actinobacillus pleuropneumoniae* (Haemophilus pleuropneumoniae).  
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 CC *Actinobacillus*.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S 4074 / SEROTYPE 1;  
 RX MEDLINE=91348845; PubMed=1879928;  
 RA Frey J., Meier R., Gyg D., Nicolet J.;  
 RT "Nucleotide sequence of the hemolysin I gene from *Actinobacillus*  
 RT *pleuropneumoniae*.";  
 RL Infect. Immun. 59:3026-3032(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S 4074 / SEROTYPE 1;  
 RX MEDLINE=94237497; PubMed=8181764;  
 RA Frey J., Haldemann A., Nicolet J., Boffini A., Prentki P.;  
 RT "Sequence analysis and transcription of the apxi operon (hemolysin I)  
 RT from *Actinobacillus pleuropneumoniae*.";  
 RL Gene 142:97-102(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ISOLATE CVI 13261 / SEROTYPE 9;  
 RX MEDLINE=93366425; PubMed=8359891;  
 RA Jansen R., Briatore J., Kamp E.M., Gielkens A.L.J., Smits M.A.;  
 RT "Structural analysis of the *Actinobacillus pleuropneumoniae* RTX-toxin  
 RT I (ApxI) operon.";  
 RL Infect. Immun. 61:3688-3695(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S 4074 / SEROTYPE 1;  
 RX Chang Y., Wang Y., Chin N.;  
 RL submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ONE OF THE VIRULENCE FACTORS OF A PLEUROPNEUMONIAE,  
 CC WHICH HAS A STRONG HEMOLYTIC ACTIVITY AND IS CYTOTOXIC FOR  
 CC ALVEOLAR MACROPHAGES AND NEUTROPHILS.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING  
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC

[illegible]

RESULT	11
RT12.ACPEL	STANDARD; PRT; 1023 AA.
ID RT12.ACPEL	PRT; 1023 AA.
AC P55129;	
DT 01-OCT-1996 (Rel. 34, Created)	
DT 01-OCT-1996 (Rel. 34, Last sequence update)	
DT 30-MAY-2000 (Rel. 39, Last annotation update)	
DE RYX-I TOXIN DETERMINANT A FROM SEROTYPES 5/10 (APX-IA) (HEMOLYSIN IA)	
DE (HLV-IA) (CTYOLYSIN IA) (CLY-IA).	
GN APXIA OR CLYIA OR HLVIA.	
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).	
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;	
OC Actinobacillus.	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=13039 / SEROTYPE 10;	
RX MEDLINE=94276858; PubMed=8007819;	
RA Nagai S., Yagihashi T., Ishihama A.;	
RT "DNA sequence analysis of an allelic variant of the Actinobacillus	
RT pleuropneumoniae-RTX-toxin I (ApXia) from serotype 10.";	
RL Microb. Pathog. 15:485-495(1993).	
RN [2]	
RP SEQUENCE FROM N.A.	
RC STRAIN=K17 / SEROTYPE 5;	
RX MEDLINE=96401417; PubMed=8807793;	
RA Chin N., Frey J., Chang C.F., Chang Y.F.;	
RT "Identification of a locus involved in the utilization of iron by	
RT Actinobacillus pleuropneumoniae.";	
RL FEMS Microbiol. Lett. 143:1-6(1996).	
RN [3]	
RP SEQUENCE OF 886-1023 FROM N.A.	
RC STRAIN=K17 / SEROTYPE 5;	
RX MEDLINE=93366425; PubMed=8359891;	
RA Jansen R., Briatore J., Kamp E.M., Gielkens A.L.J., Smits M.A.;	
RT "Structural analysis of the Actinobacillus pleuropneumoniae-RTX toxin	
RT I (ApX1) operon.";	
RL Infect. Immun. 61:3688-3695(1993).	
CC -1- FUNCTION: ONE OF THE VIRULENCE FACTORS OF A. PLEUROPNEUMONIAE,	
CC WHICH HAS A STRONG HEMOLYTIC ACTIVITY AND IS CYTOTOXIC FOR	
CC ALVEOLAR MACROPHAGES AND NEUTROPHILS.	
CC -1- SUBCELLULAR LOCATION: SECRETED.	
CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING	
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC	
CC ACTIVITY.	
CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE	
CC INVOLVED IN PORE FORMATION BY APXIC. THE CYTOTOXIN (BY SIMILARITY).	
CC -1- PTM: PALMITOYLATED BY APXIC. THE TOXIN ONLY BECOMES ACTIVE WHEN	
CC MODIFIED (BY SIMILARITY).	
CC -1- MISCELLANEOUS: APXIA IS PARTIALLY DELETED IN SEROTYPES 2, 4, 6, 7	

8, 12, AND TOTALLY DELETED IN SEROTYPE 3.  
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 10.  
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D16582; BAA04014.1; -  
 DR EMBL: U04954; AAB17220.1; -  
 DR EMBL: X73116; CAA51546.1; -  
 DR INTERPRO: IPR001343; -  
 DR PIR: PF00353; hemolysinCbind; 2.  
 DR PRINTS: PR00313; CABINDNGRPT.  
 DR PROSITE: PS00330; HEMOLYSIN\_CALCIUM; 2.  
 KM Hemolysin; Toxin; Cytolysin; Cytotoxin; Repeat; Calcium;  
 KW Transmembrane; Lipoprotein; Palmitate.  
 FT TRANSMEM 226 256 POTENTIAL.  
 FT TRANSMEM 297 326 POTENTIAL.  
 FT TRANSMEM 367 406 POTENTIAL.  
 FT DOMAIN 722 845 13 X REPEATS, GLY-RICH.  
 FT REPEAT 722 727 1.  
 FT REPEAT 731 736 2.  
 FT REPEAT 740 745 3.  
 FT REPEAT 749 754 4.  
 FT REPEAT 758 763 5.  
 FT REPEAT 767 772 6.  
 FT REPEAT 776 781 7.  
 FT REPEAT 785 790 8.  
 FT REPEAT 794 799 9.  
 FT REPEAT 813 818 10.  
 FT REPEAT 822 827 11.  
 FT REPEAT 831 836 12.  
 FT REPEAT 840 845 13.  
 FT CONFLICT 210 217 AMPYLITLA -> GNAISNTR (IN REF. 2).  
 FT CONFLICT 581 581 E -> Q (IN REF. 2).  
 FT CONFLICT 687 688 TC -> R (IN REF. 2).  
 FT CONFLICT 1015 1015 F -> L (IN REF. 2).  
 SQ SEQUENCE 1023 AA; 110129 MW; 183C7C15EE57DB55 CRC64;

Query Match 23.1%; Score 828.5; DB 1; Length 1023;  
 Best Local Similarity 36.7%; Pred. No. 4.3e-34;  
 Matches 213; Conservative 91; Mismatches 187; Indels 89; Gaps 17;

DB 85 SQHMEYGLRPGSGSDMSYGLRPGSSFPKTKAKKIIILYIPQNYQYDTEGCGLODLYKA 144  
 17 NQHTSKAASGGA-----LKNGLGVKQAG-QKILILYIPDYAST--GSLNDLYKA 67  
 145 AEEIGIEVQREERNIATAQTSLGTQTAIGLTERGIYLSAPQIDKLQKT-KAGQAL-G 202  
 68 AEAIGIEVHRSKNTALAKELFTTEKILGSESGIALFAQFPKLNKNKLSLGG 127  
 203 SAESTIVQANKAKIVLSGIQSLISGLAGMDLDEALQNNNSQH-----ALAKAGLELNS 257  
 128 SSEALGQRLKQTALSLAQSLFGLTAIGMDLSLRRRRNCEDEVSGSELAKAGVDLAAQ 187  
 DB 258 LIENIANSVKTLDEGEQISQSGSKLQNKIGLGTCDKRIKNGIGDKAGLGIDVYSGLS 317  
 188 LVDNLTASATGIVDAEDQGLAMPYLYL-ALSGLASKLNNLPDLISLGGPGDAVSGILS 246  
 318 GATALLVADKNASTAKKVGAFELANQVGNITKAVSSYIIAQRVAGLSTGPPVALI 377  
 247 VVSASFILSNKDDADGTRAAAGIEISTKILNIGKAVSYIIIAQRVAGLSTTAATGILI 306  
 DB 378 ASTVSLAIPLAFAIGIADKFNHAKSLSEYAEERFKLGYDGNLLAEYQGTIDASYTA 437  
 307 GSVVALAISPLISFLNVADFERAKQLEQYSERFKKFGEYEGDLSLASFRETCAIEALTT 366

DB 438 INTALAAAGVSAANA-----DLFEKYVHNIVTNS 470  
 367 INSVLASASAGGAATGSLVAPYALYSATITGIISGLDASKAKAIFERVATKLANMD 426  
 DB 471 KREKVTIONMF-----READFAKE---VPNYKATKDEKIEEIGONGER----- 511  
 427 EMEKHKKNGYFENGVDARHSAPLEDTFELLSQY--NKEYSVERVAIIQOQPDVINGELA 484  
 DB 512 -ITSQVVDLILKG-----NKRITODELSKYVDVYELKSKNNTNSIDKLI--SSVSA 562  
 485 GTRRGGAARSKGKAVYDFEEEGKL--LEKDPDRFD---KKVFDPLEGRIDLSINK 535  
 DB 563 FTSSNDSRNVLYAPTSMDQSLSSIQ-----FAGSGQHW 596  
 536 TTLKFTPTVPFVTAAGEIERKQTKGYEWTFLFVAGKEW 575

RESULT 12  
 ID HLY1\_ECOLI STANDARD: PRT: 1023 AA.  
 AC P09983;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HEMOLYSIN, CHROMOSOMAL.  
 GN HLYA.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-196 / SEROTYPE O4;  
 RX MEDLINE-85234404; PubMed-3891743;  
 RA Fellee T., Pellett S., Welch R.A.;  
 RT "Nucleotide sequence of an Escherichia coli chromosomal hemolysin";  
 RL J. Bacteriol. 163:94-105(1985).  
 RN (2)  
 RP SEQUENCE OF 1-44 FROM N.A.  
 RC STRAIN-2001;  
 RX MEDLINE-85258115; PubMed-3894051;  
 RA Nicaud J.-M., Mackman N., Gray L., Holland I.B.;  
 RT "Characterisation of HLYC and mechanism of activation and secretion  
 of haemolysin from E. coli 2001";  
 RL FEBS Lett. 187:339-344(1985).  
 CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD  
 CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY  
 CC DEFINED.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING  
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC  
 CC ACTIVITY.  
 CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE  
 CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN.  
 CC -1- PTM: PALMITOYLATED BY HLYC. THE TOXIN ONLY BECOMES ACTIVE WHEN  
 CC MODIFIED.  
 CC -1- DISEASE: THE HEMOLYSIN OF E. COLI IS PRODUCED PREDOMINANTLY BY  
 CC STRAINS CAUSING EXTRAIESTINAL INFECTIONS, SUCH AS THOSE OF THE  
 CC URINARY TRACT.  
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M10133; AAA23975.1; -  
 DR EMBL: X02768; CAA26546.1; -  
 DR PIR: A24433; LEECA.  
 DR INTERPRO: IPR001343; -  
 DR PIR: PF00353; hemolysinCbind; 2.









Matches: 200; Conservative 87; Mismatches 128; Indels 106; Gaps 13;

114 KTGAKKIITYIPONYGYDEQNGLODLVKAABELGEYOREERNINATPQSTIGTOTA 173  
 44 KTG-KKLTLYIKPNY-----KKNGLTALLIKAAOKLGEYHEKODPALNGLNGLTKL 98  
 174 IGTERTGIVLSAFOIDKLQKTR-AQOALGSASISIVOMANKAKTVLSIGSISLVACM 232  
 99 LGTERTGLTFAPELDKNLQNGKHLNSVSGSTGNLTKAIDKQVSGTIGLOAFILNTAFSGM 158  
 233 DLD---EAQONNSN--QHAKRAGLELNSLINIANSVKTDDEFEQISQFSKQONIK 287  
 159 DLDALIKRQNGKNTVDYOLAKASLNLINELIGTISSTNNVTFFSKQLKKLEALGOYK 218  
 288 GLGTLDLKNIGKIDKAGLGDVIGSLSGATPAALVLADKNASTAKKVGAFELANOV 347  
 219 HGSEFEDKLKNLPKLGNLGKGYALSGLVSAISALLANKHADTKAAMAAELNKKVL 278  
 348 GNTKAVSSYIIAQRYAAGISTGPVAAALASTVSLAISPLAFAGIADKENHAKSLESYA 407  
 279 GNIGKAITOYLINQRAAAGISTGTPVAGIASVSLAISPLSLGIAKQDRARMLEAYS 338  
 408 ERKKTIGYGDNLIAEYOKGTGIDASVTAINTALAIAGVSA-----453  
 339 KRKKKGYNGDSLGLQFYKNYGIADAITITNTVLSAIAAGVAGASGLVGAPIGLVS 398  
 454 -----ADLFEKVK-----HNLVIT 468  
 399 AITSLISGLIDASKOAVSEIHANQADKIKAMENKYGKNTSENGYDARHSAPLEDISKLP 458  
 469 NSRKEK-----VTIOWMFREADPAKEVYKATKDEKIEIIG--ONGERITSKO-- 516  
 459 NELREKYKTEINISITQOGW-----DQRIEGLAGITRNGDRIGSKAV 501  
 517 VDDL-----IAKNGKITODELSKVVDNEL--LHKSANT 550  
 502 VDYLKKEELAKHSDKFTQIOLDIPKINIDLGSIGKSTILT 542

RESULT 15  
 CYAA-BORPE  
 ID CYAA-BORPE STANDARD; PRT: 1706 AA.  
 AC P15318;  
 DT 01-APR-1990 (rel. 14, Created)  
 DT 01-APR-1990 (rel. 14, last sequence update)  
 DT 30-MAY-2000 (rel. 39, last annotation update)  
 DE BIFUNCTIONAL HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (CYCLOLYSIN) (ACT)  
 DE (AC-HLY) [CONTRAINS: CALMODULIN-SENSITIVE ADENYLATE CYCLASE  
 DE (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (ADENYL CYCLASE); HEMOLYSIN].  
 GN CYA OR CYAA.  
 OS Bordetella pertussis.  
 OC Bordetella pertussis; beta subdivision; Alcaligenaceae;  
 OC Bordetella.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-18323;  
 RX MEDLINE-88216178; PubMed-2897067;  
 RA Glaser P., Ladant O., Sezer O., Pichot F., Ullmann A., Danchin A.;  
 RT "The calmodulin-sensitive adenylate cyclase of Bordetella pertussis:  
 RT cloning and expression in Escherichia coli.";  
 RT Mol. Microbiol. 2:19-30(1988).  
 RL (2)  
 RP SEQUENCE OF 1489-1706 FROM N.A., AND BIFUNCTIONAL PROTEIN DESCRIPTION.  
 RC STRAIN-18323;  
 RX MEDLINE-89091151; PubMed-2905265;  
 RA Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;  
 RT "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-  
 RT haemolysin bifunctional protein of Bordetella pertussis.";  
 RL EMBO J. 7:3997-4004(1988).  
 RN (3)  
 RP DOMAINS  
 RX MEDLINE-91177021; PubMed-2007407;  
 RA Munier H., Gilles A.-M., Glaser P., Danchin A., Sarfati R., Barzu O.;

RT "Isolation and characterization of catalytic and calmodulin-binding  
 RT domains of Bordetella pertussis adenylate cyclase.";  
 RL Eur. J. Biochem. 196:469-474(1991).  
 RN (4)  
 RP MUTAGENESIS.  
 RX MEDLINE-89251630; PubMed-2542030;  
 RA Glaser P., Elmaouglou-Lazaridou A., Krin E., Ladant D., Barzu O.,  
 RA Danchin A.;  
 RT "Identification of residues essential for catalysis and binding of  
 RT calmodulin in Bordetella pertussis adenylate cyclase by site-directed  
 RT mutagenesis.";  
 RL EMBO J. 8:967-972(1989).  
 RN (5)  
 RP MUTAGENESIS.  
 RX MEDLINE-91266896; PubMed-2050107;  
 RA Glaser P., Munier H., Gilles A.-M., Krin E., Porumb T., Barzu O.,  
 RA Sarfati R., Pelleuer C., Danchin A.;  
 RT "Functional consequences of single amino acid substitutions in  
 RT calmodulin-activated adenylate cyclase of Bordetella pertussis.";  
 RL EMBO J. 10:1683-1688(1991).  
 RN (6)  
 RP REVIEW.  
 RX MEDLINE-93119764; PubMed-8418825;  
 RA Danchin A.;  
 RT "Phylogeny of adenylate cyclases.";  
 RL Adv. Second Messenger Phosphoprotein Res. 27:109-162(1993).  
 RN (7)  
 RP PALMITOYLATION AT LYS-983.  
 RX MEDLINE-95025937; PubMed-7939682;  
 RA Hackett M., Guo L., Shabunowitz J., Hunt D.F., Hewlett E.L.;  
 RT "Internal lysine palmitoylation in adenylate cyclase toxin from  
 RT Bordetella pertussis.";  
 RL Science 266:433-435(1994).  
 RN (8)  
 RP PALMITOYLATION AT LYS-860.  
 RX MEDLINE-9214144; PubMed-10196151;  
 RA Basar T., Havlicek V., Bezouskova S., Halada P., Hackett M., Sebo P.;  
 RT "The conserved lysine 860 in the additional fatty-acylation site of  
 RT Bordetella pertussis adenylate cyclase is crucial for toxin function  
 RT independently of its acylation status.";  
 RL J. Biol. Chem. 274:10777-10783(1999).  
 CC -1- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF  
 CC BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN  
 CC CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL  
 CC CELL FUNCTION.  
 CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.  
 CC -1- ENZYME REGULATION: ACTIVATED BY HOST CALMODULIN.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING  
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC  
 CC ACTIVITY (BY SIMILARITY).  
 CC -1- PTM: RELEASED IN A PROCESSED FORM.  
 CC -1- PTM: PALMITOYLATED BY CYAC. THE TOXIN ONLY BECOMES ACTIVE WHEN  
 CC MODIFIED IN POSITION LYS-983.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO ADENYLATE CYCLASE  
 CC CLASS-2 FAMILY.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE RTX  
 CC PROKARYOTIC TOXIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: Y00545; CAA68613.1; -  
 CC EMBL: X14199; CAA32411.1; -  
 CC EMBL: A07292; CAA00653.1; -  
 CC EMBL: A14850; CAA01202.1; -  
 CC PIR: S00893; CYBC.  
 CC PIR: S09403; S09403.

```

DR INTERPRO: IPR001343;
DR PFAM: PF00353; hemolysinCbind; 5.
DR PRINTS: PR00313; CABINDNGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCTIM; 5.
KW Lyase; CAMP synthesis; ATP-binding; Hemolysis; Toxin; Virulence;
KW Whooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
FT CHAIN 1 312 CALMODULIN-SENSITIVE ADENYLATE
FT CHAIN 1 312 CYCLASE.
FT CHAIN 313 1706 HEMOLYSIN (BY SIMILARITY TO E.COLI
FT 1 HEMOLYSIN HLVA).
FT DOMAIN 1 399 A, CATALYTIC.
FT DOMAIN 400 912 B, ALA/GLY-RICH.
FT DOMAIN 913 1656 C.
FT DOMAIN 1657 1706 D, ASP/GLY-RICH.
FT NP_BIND 349 356 ATP (POTENTIAL).
FT DOMAIN 913 1610 28 X REPEATS, GLY-RICH.
FT REPEAT 1015 1020 1.
FT REPEAT 1024 1029 2.
FT REPEAT 1033 1038 3.
FT REPEAT 1042 1047 4.
FT REPEAT 1051 1056 5.
FT REPEAT 1060 1065 6.
FT REPEAT 1080 1085 7.
FT REPEAT 1165 1170 8.
FT REPEAT 1174 1179 9.
FT REPEAT 1183 1188 10.
FT REPEAT 1203 1208 11.
FT REPEAT 1280 1285 12.
FT REPEAT 1289 1294 13.
FT REPEAT 1298 1303 14.
FT REPEAT 1307 1312 15.
FT REPEAT 1316 1321 16.
FT REPEAT 1325 1330 17.
FT REPEAT 1345 1350 18.
FT REPEAT 1421 1426 19.
FT REPEAT 1430 1435 20.
FT REPEAT 1439 1444 21.
FT REPEAT 1448 1453 22.
FT REPEAT 1456 1461 23.
FT REPEAT 1565 1570 24.
FT REPEAT 1574 1579 25.
FT REPEAT 1583 1588 26.
FT REPEAT 1593 1598 27.
FT REPEAT 1605 1610 28.
FT LIPID 860 860 PALMITATE.
FT LIPID 983 983 PALMITATE.
FT MUTAGEN 188 188 D->E,N,Y,H: LOSS OF ACTIVITY.
FT MUTAGEN 190 190 D->N,Y,H: LOSS OF ACTIVITY.
FT MUTAGEN 298 298 H->R,P,L: LOSS OF ACTIVITY.
FT MUTAGEN 301 301 E->Q,K: LOSS OF ACTIVITY.
SQ SEQUENCE 1706 AA; 177506 MW; F00744524BD442E CRC64;

```

Query Match 5.9%; Score 212; DB 1; Length 1706;

Best Local Similarity 22.3%; Pred. No. 0.003; Indels 94; Gaps 21;

Matches 114; Conservative 84; Mismatches 220; Indels 94; Gaps 21;

```

QY 89 SYGLRPGSGSODWSYGLRPGSSFPKTKAKIILYIPONYQYDTQNGIOLDLVKAAREL 148
DB 349 AYGV-AGKSLFDGAGAPGVSSGRKSPDLVTPAS-----PGLR-----RPSL 394
QY 149 GIEVOREER-----NNIATAOTSIGTIQTAIGLTERGIIVLSAPQIDKLIQTKAGQALGS 203
DB 395 G-AVERODSGYSDLDGVSRSFSLEVSDM-----AAVEAAELEMTROYVLHAGARQDD 446
QY 204 AESTIVQNMK--AKTVLGSIGSILGSVLGMDLEALQNNNSQHALLA-----KAGLELT 255
DB 447 AEPGVSGASAHMGORALQAGQAVANA-----QRLVHAIALMTQFGRAGSTNT 493
QY 256 NSLIENIANSVKTLDEFGQLSQFGSKLQNIKGLTGLDKLKNIGLKDAGLGLDVISGL 315
DB 494 PQEASLSLAAYVGLGEASSAVAETVS-----GFFRSSRWAGSGFVAGGAMALGGGI 545

```

```

QY 316 LSGATAVLVLADKNKSTAKYVAGFELANOVYGNITKAVSYIL-----AQRYAAGLSSTG 371
DB 546 AAVAGAGMSLTD-DAPAGQKAAAGAEIALQLTGTVELASSITALAARGVTSGLQVAG 604
QY 372 PYAALIASTVSLAISPFAFIADKFNNHAKSLESYAERPKKLGYGDNLLAEYQGTGTI 431
DB 605 ASAGAAAGALAAALSPMETIYGLVQOSHADQDLKLAQESSAYGEGDALIAQLYRDKTAA 664
QY 432 DASVTAINTAALAAIAGVS-AAAADLTFEYKHNLVITNSKREKVT-----IOMNFREA- 484
DB 665 ECAVAGVSAYLSTVGAAVSIAAASVGVAPV---AVVTSLLTGALNGILRGVQPIIEKL 721
QY 485 --DFAKEVPN-----YKATDEKIEELIGNGERITRSQYDDLIANGNK-----ITQD 531
DB 722 ANDYARKIDELGGPOAYFEKNLQARHEQLANSDDLRL---KMLADLQAGMNASSVIGVOTT 778
QY 532 ELKVVYDNYELLKHSKNVTNSLDKLISSVSAP 563
DB 779 EISK--SALELTAITGMADN-----LKSVDVF 803

```

Search completed: March 2, 2001, 11:01:34  
Job time: 449 sec

**THIS PAGE BLANK (USPTO)**

OM of: US-09-306-689-12 to: SPTREMBL\_15:\* out\_format : pfs  
 Date: Mar 2, 2001 10:49 AM

About: Results were produced by the Gencore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frmet-n2p model -DEV=rlp  
 -Q/cgn2\_1/USPTO.spool/US09306689/runat.02032001\_102828\_9616/app\_query.fasta.1.2389  
 -DB=SPTREMBL\_15 -QFMT=fasta -SUFFIX=rspt -GAPOP=12.000  
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
 -GAPOP=4.000 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
 -FCAPOP=6.000 -FCAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bloms62  
 -TRANS=human0.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09306689\_@CGN1\_1\_639 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
 -MAIT -THREADS=1

## Search information block:

Query: US-09-306-689-12  
 Query length: 2088  
 Database: SPTREMBL\_15:\*  
 Database sequences: 374700  
 Database length: 117207915  
 Search time (sec): 337.480000

## score\_list:

Sequence	Strd	Orig	ZScore	EScore	len	Documentation
SP_bacteria:068403	+	377.50	366.99	1.2e-12	181	068403 escherichia coli. alpha
SP_bacteria:068404	+	377.50	366.99	1.2e-12	181	068404 escherichia coli. alpha
SP_bacteria:070070	+	372.50	364.28	2.3e-12	181	070070 escherichia coli. alpha
SP_invertebrate:09617	+	208.50	191.73	0.0010	1729	09617 drosophila melanogaster
SP_bacteria:091469	+	205.50	189.00	0.0014	1706	091469 bordetella parapertussis
SP_bacteria:096333	+	199.00	177.44	0.0031	3381	096333 streptococcus cristatus
SP_bacteria:052781	+	190.50	178.27	0.0085	1112	052781 campylobacter fetus. su
SP_bacteria:061869	+	187.50	179.04	0.0121	707	061869 mus musculus (mouse). ke
SP_bacteria:091964	+	187.50	176.93	0.0121	922	091964 moraxella catarrhalis. v
SP_invertebrate:09766	+	185.50	174.09	0.0155	1039	09766 drosophila melanogaster
SP_invertebrate:013083	+	173.00	165.20	0.0697	722	013083 escherichia coli. flagel
SP_bacteria:098073	+	170.00	164.32	0.0997	565	098073 escherichia coli. flagel
SP_bacteria:098073	+	169.00	155.79	0.1147	1467	098073 deinetococcus radiodurans
SP_bacteria:098073	+	168.50	156.99	0.1213	1190	098073 xylella fastidiosa. su
SP_bacteria:068729	+	166.50	153.13	0.1553	1525	068729 yersinia pestis. prote
SP_bacteria:098729	+	166.00	158.24	0.1627	756	098729 deinetococcus radiodurans
SP_bacteria:053505	+	165.00	154.25	0.1850	1109	053505 campylobacter fetus. ty
SP_invertebrate:097054	+	165.00	147.43	0.1882	2614	097054 dictyostellum disco
SP_bacteria:047226	+	164.00	158.67	0.2060	565	047226 escherichia coli. flagel
SP_invertebrate:080179	+	163.00	149.91	0.2371	1510	080179 streptococcus thermoph
SP_invertebrate:09NHX4	+	162.00	150.02	0.2668	1323	09NHX4 drosophila melanogaster
SP_fungi:013936	+	161.00	152.25	0.2994	990	013936 schizosaccharomyces pom
SP_rodent:09MYR8	+	160.50	152.25	0.3170	836	09MYR8 rattus norvegicus (rat)
SP_bacteria:09PD50	+	160.00	144.61	0.3428	2059	09PD50 xylella fastidiosa. su
SP_bacteria:034071	+	159.00	146.10	0.3845	1517	034071 streptococcus thermoph
SP_human:09N202	+	159.00	142.00	0.3885	2540	09N202 homo sapiens (human)
SP_bacteria:046037	+	158.50	149.47	0.4046	936	046037 campylobacter fetus. sag
SP_phage:09T1A7	+	158.50	144.30	0.4099	1794	09T1A7 bacteriophage al18. put
SP_bacteria:092544	+	158.00	147.32	0.4316	1156	092544 streptomyces coelicolor
SP_bacteria:091A90	+	158.00	146.22	0.4328	1328	091A90 xanthomonas campestris
SP_bacteria:054415	+	157.50	153.40	0.4511	507	054415 salmonella montevideo. ph
SP_mammal:097643	+	157.00	143.67	0.4784	462	097643 lama glama (llama). fib
SP_bacteria:098X38	+	156.00	147.49	0.5469	892	098X38 moraxella catarrhalis. v
SP_bacteria:054515	+	155.50	151.50	0.5746	508	054515 salmonella pensacola. ph
SP_bacteria:09JW98	+	155.50	145.59	0.5831	1068	09JW98 neisseria meningitidis

seq\_name: sp\_bacteria:068403  
 seq\_documentation\_block:  
 ID 068403 PRELIMINARY; PRT: 181 AA.

AC 068403:  
 DT 01-AUG-1998 (TREMREL. 07, Created)  
 DT 01-AUG-1998 (TREMREL. 07, Last sequence update)  
 DT 01-MAY-2000 (TREMREL. 13, Last annotation update)  
 DE ALPHA HEMOLYSIN (FRAGMENT).  
 GN hylA.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 NX NCBI\_TaxID=562;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ECOR 24;  
 RX MEDLINE=98155142; PubMed=9495754;  
 RA Boyd E.F., Hartl D.L.;  
 RT "Chromosomal regions specific to pathogenic isolates of Escherichia  
 coli have a phylogenetically clustered distribution".  
 RL J. Bacteriol. 180:1159-1165(1998).  
 DR EMBL: AF037572; AAC38261.1; -  
 DR INTERPRO: IPR000911; -  
 DR PROSITE: PS00359; RIBOSOMAL\_L11; UNKNOWN\_1.  
 FT NON\_TER 1  
 FT SEQUENCE 181 AA; 19071 MW; 570928D2233B4C04 CRC64;

alignment\_scores:  
 Quality: 377.50 Length: 181  
 Ratio: 2.640 Gaps: 3  
 Percent similarity: 79.006 Percent identity: 45.856

alignment\_block:  
 US-09-306-689-12 x 068403 ..

Align seg 1/1 to: 068403 from: 1 to: 181

403 GGTATGCTTACAGATTATGACCAAGCGCCGAGAGTGGCATTTGA  
 1 GlySerSerLeuAsnSplLeuValArgThrAlaAspGluLeuGlyIleG 17  
 453 GGTCAAGAGAGAGAGCAATATATTCACACAGCTCAACCCATTAG  
 17 ValGlnTyrAspGluLysAsnGlyThrAlaThrLysGlnValAlpHeG 34  
 503 GCAAGATTCAACCGCTATTGCTTACAGAGCGTGGCATTTGTTATCC  
 34 LThrAlaGluLysLeuIleGlyLeuThrGluArgGlyValAlaIleHe 50  
 553 GCTCCAAATGATTAATTTGTCACAGAAA...ACTAAAGCAGGCCAACG  
 51 AlaProGlnLeuAspLysLeuGlnLysTyrGlnLysAlaGlyAsnLys 67  
 600 ATTA...GGTTGCCGAAGCAATGTGTACAAATTAAGCAATTAAGCCAAA  
 67 sleuGlyGlySerAlaGluAsnIleGlyAspAsnLeuGlyLysAlaGlyS 84  
 647 CAGTATTAATTCGATTAATTAATTTAGCGTCAATATTTGCTGGAAG  
 84 ervAlaLeuSerThrPheGlnAsnIlePheLeuGlyAlaAlaLeuSerMet 100  
 697 GATTATGAGAGCGCTTACAG.....ATAACAGCAACCA 731  
 101 LysIleAspGluLeuIleLysGlnLysSerGlyAsnValSerS 117  
 732 ACATGCTTGTCTTAAGCTGGCTTGAGCTTAACAAATTCATTAATGA 781  
 117 rSerGluLeuAlaLysAlaSerIleGluLeuIleAsnGlnLeuValAspT 134  
 782 AATTTGCTTAATTCAGTAATAACACTGACGAATTTGCTGACCAATTA 831  
 134 hValAlaSerIleAsnAsnValAsnSerPheSerIleGlnLeuVal 150  
 832 CAATTTGCTCAAAACTCAAAATATCAAGGCTTACGACCTTACGAGA 881

151 LysleuGlySerValIeuSerAsnThrLysHisLeuAsnGlyValGlyAs 167  
 882 CAAACTCAAAAATATCGTGCTGATTAAGCTGCTTGT 924  
 :|||||:|||||: :|||||: ||| |||  
 167 nLysleuGlnAsnLeuProAsnLeuAspAsnIleGlyAlaGly 181  
 seq\_name: sp\_bacteria:068404

seq\_documentation\_block:  
 ID 068404 PRELIMINARY: PRT; 181 AA.  
 AC 068404;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)  
 DE ALPHA HEMOLYSIN (FRAGMENT).  
 GN HYL.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ECOR 43;  
 RX MEDLINE=98155142; PubMed=9495754;  
 RA Boyd E.F., Hartl D.L.;  
 RT "Chromosomal regions specific to pathogenic isolates of Escherichia  
 coli have a phylogenetically clustered distribution."  
 RL J. Bacteriol. 180:1159-1165(1998).  
 DR EMBL: AF037573; AAC38262.1; -;  
 DR INTERPRO: IPR000911; -;  
 DR PROSITE: PS00359; RIBOSOMAL\_L11; UNKNOWN\_1.  
 FT NON\_TER 1  
 SQ SEQUENCE 181 AA; 19101 MW; A3586DC663B571E CRC64;

alignment\_scores:  
 Quality: 377.50 Length: 181  
 Ratio: 2.640 Gaps: 3  
 Percent Similarity: 79.006 Percent Identity: 45.856

alignment\_block:  
 US-09-306-689-12 x 068404 ..

Align seg 1/1 to: 068404 from: 1 to: 181

```

403 GGTATGTTTACAGATTGATTCAGACCGCGCAAGATTGGGATTGA 452
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 GlySerSerLeuAsnAspLeuValArgThrAlaAspGluLeuGlyIleG1 17
453 GGTACAAAGAGAGACGCAATTAATTGCAACAGCTCAACACGATTAG 502
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 uValGlnTyAspGluLysAsnGlyThrAlaIleThrLysGlnValPheG 34
503 GCACGATTCAAAACCGCTATTGGCTTAACCTGAGCGTGGCATTGTTC 552
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 LyrThrAlaGluLysLeuIleGlyLeuThrGluArgGlyValAlaIlePhe 50
553 GCTCCACAATTAATTGCTACAGAAA...ACTAAAGCAGGCCAAGC 599
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 AlaProGlnLeuAspLysLeuGlnLysTyGlnLysAlaGlyAsnLys 67
600 ATTA...GGTTCGCCGAAAGCATTTGTAACAATGCAATTAAGCCAAA 646
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 slenGlyGlySerAlaGluAsnIleGlyAspAsnLeuGlyLysAlaGlyS 84
647 CTGATTAATCTGCATTCATCTATTTAGCGTCAAGTATGGCTGAATG 696
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84 erValIeuSerThrPheGlnAsnPheLeuGlyThrAlaLeuSerSerMet 100
697 GATTAGTAGGCGCTTACAG.....ATAACAGCAACCA 731
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 LysIleAspGlnLeuIleLysLysGlnLysSerGlnAsnValSerSe 117

```

732 ACATGCTCTTGCTAAAGCTGGCTTGAGCTAACAATTCATTAATGAAA 781  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 117 rSerGlnLeuAlaLysAlaSerIleGlyLeuIleAsnGlnLeuValAspT 134  
 782 ATATTGCTAATTCGTAATAAACTTGACGCAATTTGGTGAGCAATTAAGT 831  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 134 hrAlaAlaSerIleAsnAsnValAsnSerPheSerGlnGlnLeuAsn 150  
 832 CAATTGTTTCAAAATACAAATATATCAAGCGCTAGGACTTAAAGAGA 881  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 151 LysleuGlySerValIeuSerAsnThrLysHisLeuAsnGlyValGlyAs 167  
 882 CAAACTCAAAAATATCGTGCTGATTAAGCTGCTTGT 924  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 167 nLysleuGlnAsnLeuProAsnLeuAspAsnIleGlyAlaGly 181  
 seq\_name: sp\_bacteria:070070

seq\_documentation\_block:  
 ID 070070 PRELIMINARY: PRT; 181 AA.  
 AC 070070;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)  
 DE ALPHA HEMOLYSIN (FRAGMENT).  
 GN HYL.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ECOR 48, ECOR 51, ECOR 52, ECOR 54, ECOR 60, ECOR 63;  
 RX MEDLINE=98155142; PubMed=9495754;  
 RA Boyd E.F., Hartl D.L.;  
 RT "Chromosomal regions specific to pathogenic isolates of Escherichia  
 coli have a phylogenetically clustered distribution."  
 RL J. Bacteriol. 180:1159-1165(1998).  
 DR EMBL: AF037579; AAC38268.1; -;  
 DR EMBL: AF037574; AAC38263.1; -;  
 DR EMBL: AF037575; AAC38264.1; -;  
 DR EMBL: AF037576; AAC38265.1; -;  
 DR EMBL: AF037577; AAC38266.1; -;  
 DR EMBL: AF037578; AAC38267.1; -;  
 DR INTERPRO: IPR000255; -;  
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN\_1.  
 DR PROSITE: PS00359; RIBOSOMAL\_L11; UNKNOWN\_1.  
 FT NON\_TER 1  
 SQ SEQUENCE 181 AA; 19104 MW; 61716DD67DD6D613 CRC64;

alignment\_scores:  
 Quality: 372.50 Length: 181  
 Ratio: 2.605 Gaps: 3  
 Percent Similarity: 79.006 Percent Identity: 45.304

alignment\_block:  
 US-09-306-689-12 x 070070 ..

Align seg 1/1 to: 070070 from: 1 to: 181

```

403 GGTATGTTTACAGATTGATTCAGACCGCGCAAGATTGGGATTGA 452
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 GlySerSerLeuAsnAspLeuValArgThrAlaAspGluLeuGlyIleG1 17
453 GGTACAAAGAGAGACGCAATTAATTGCAACAGCTCAACACGATTAG 502
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 uValGlnTyAspGluLysAsnGlyThrAlaIleThrLysGlnValPheG 34
503 GCACGATTCAAAACCGCTATTGGCTTAACCTGAGCGTGGCATTGTTC 552
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 LyrThrAlaGluLysLeuIleGlyLeuThrGluArgGlyValThrIlePhe 50

```



```

828 TAGCTAATTGTTCAAACTACAAATATCAAGGCTTAGGACT... 873
    |||::: ||| |||||
693 oGlyIleGlyGlnThrGlyProGlyTYrGlyThrGlnP 710
    |||::: ||| |||||
874 .....TTAGAGACAAACTC..... 888
    |||::: ||| |||||
710 roGlyValGlyThrGlnThrGlyThrGlyProGlyTYrGlyAlaGln 726
    |||::: ||| |||||
889 AAAAATTCGGTGCTGATTAAGGCTGGCTT..... 921
    |||::: ||| |||||
727 ProGlyIleGlyGlnSerGlyAlaGlyGlnProGlyTYrGlySerGly 743
    |||::: ||| |||||
922 .....GGTTAGATTTATCTCAGGAC 943
    |||::: ||| |||||
743 nProGlyIleGlyGlnThrGlyGlyGlnProGlyTYrGlySerGly 760
    |||::: ||| |||||
944 TATTATCGGCGCACAGCTGCACTTACTTGCAGATTAATAATGCTTCA 993
    |||::: ||| |||||
760 InIleGlyGlyGlnThrGlyAla.....GlyGlnProSer 771
    |||::: ||| |||||
994 ACAGCTAAAGTGGGCGGTTTGAATTTGGCAAAACCAAGTTGTTGG 1043
    |||::: ||| |||||
772 TYrGlySerGlnProGlyValGlyAlaGlnAsnGlyGlyGlnProGly 788
    |||::: ||| |||||
1044 TAAATATCCAAAGCCGTT..... 1062
    |||::: ||| |||||
788 YTYrGlyTYrArgProValIleGlyGlyGlnThrGlyAlaGlyGlnProG 805
    |||::: ||| |||||
1063 .....TCTTCTTACATTTTACCCAA 1083
    |||::: ||| |||||
805 LYrGlyGlyGlnThrGlyValGlyGlySerProGlyPheLeuThrGln 821
    |||::: ||| |||||
1084 CGTCTTCAGAGGTTTATCTCAACTGGCTGGCTGCTTAAATTCG 1133
    |||::: ||| |||||
822 ProGlyIleGlyGlyIleSer.....GlyProIleGlyGlyValGly 836
    |||::: ||| |||||
1134 TTTCTACTGTTCTCTTCGATTTAGCCCATTAAGCATTTGCC..... 1173
    |||::: ||| |||||
836 YGlyGlyGlnSerGlnAlaAlaLysProGlyTYrTrpAlaGlnProGlyI 853
    |||::: ||| |||||
1174 .....GGTATTCGCGATTAATTT 1191
    |||::: ||| |||||
853 IeGlyGlyProSerArgTYrGlySerGlnProGlyIleGlyAspGlnThr 869
    |||::: ||| |||||
1192 AATCATCAAAAGTTTAGAGATTATGCCAGCTTTAAATAATTAGG 1241
    |||::: ||| |||||
870 GlyAlaIleGlnSer.....Gly 875
    |||::: ||| |||||
1242 CTATGACGGAGTATTTATTATGACAGATATCAGCGGGAACAGGACTA 1291
    |||::: ||| |||||
875 YTYrGlyGlyGlnProGlyIleSerGly...GlnThrGlyGlyGlnP 891
    |||::: ||| |||||
1292 TT.....GATGATCGGTACTGCAATTAAT..... 1317
    |||::: ||| |||||
891 roGlyTYrGlyGlyGlnAlaThrIleSerGlyLeuProGlyTYrGlyThr 907
    |||::: ||| |||||
1318 .....ACGCAATTGGCCGCTATTGCTGCTGCTGCTGCTGCTGCTG 1355
    |||::: ||| |||||
908 GlnProGlyIleGlyAlaLeuThrAlaValProGlyGly..... 920
    |||::: ||| |||||
1356 TGCAGCGGATTTAACATTTGAAAAAGTTAAACATTAATCTGTCTATCAG 1405
    |||::: ||| |||||
921 .....HistyGlyTYrGlu..... 925
    |||::: ||| |||||
1406 ATACAAAAAGAGAAAGTGACCATTCAAAACGTTCGAGAGGCTCAT 1455
    |||::: ||| |||||
926 .....ThrGlnProGlyIleGlyGlyGlnThrGlyThrAspGlnProGly 940
    |||::: ||| |||||
1456 TTTGCTAAAGAAAGTCCATATTTATTAAGCACTAAAGATGAGAAATGCA 1505
    |||::: ||| |||||
941 PheGlyGlyGln.....Pr 945
    |||::: ||| |||||
1506 AGAATATCATCGGTCAAAATGCGGAGCGATCACCTCAAGCAAGTGTGATG 1555
    |||::: ||| |||||

```

```

945 oGlyIleGlyGlnThrGlyAlaGly..... 954
    |||::: ||| |||||
1556 ATCTTATCGCAAAAGTAAGGCAAAATTAACCAAGATGAGTATCAAAA 1605
    |||::: ||| |||||
955 .....GlnProGlyTYrGlyPheIleGlyGlnProGlyIleGlyGly 968
    |||::: ||| |||||
1606 GTTGTGATTAAGTATGAAATGCTCAAAACATAGCAAAATGTGACAAACAG 1655
    |||::: ||| |||||
968 ..... 968
    |||::: ||| |||||
1656 CTTAGATAAGTAAATCTCATCTGTAAGTCAATTACTCGCTAATGATT 1705
    |||::: ||| |||||
968 ..... 968
    |||::: ||| |||||
1706 CGAANAATGTATTAGTGGCTCCAACTTCATGTTGATCAAGTTTATCT 1755
    |||::: ||| |||||
968 ..... 968
    |||::: ||| |||||
1756 TCTCTCAATTTGCTAGGAGATCTACAGATTTGAGACTACGCGCTGCGCC 1805
    |||::: ||| |||||
969 .....GlnThrGlyThrSerGlyArgGlnProGlyTYrGlyThrGlnP 983
    |||::: ||| |||||
1806 TGGCAGCGGTTCTCAGATTGAGC.....TACGCGCTGCGCTC 1843
    |||::: ||| |||||
983 oGlyIleGlyGlyGlnThrAlaAlaGlyGlnProGlyTYrGlySerGlnP 1000
    |||::: ||| |||||
1844 CGGCT.....GGCTCTAGCCAGCATTTGAGCTACGCGCTG 1878
    |||::: ||| |||||
1000 roGlyIleGlyGlyGlnThrGlyAlaGlyGlnProGlyTYrGlySer 1015
    |||::: ||| |||||
1879 CGCCCTGCGAGCGGTAGCCAA.....GATTGAGCTACGCGCT 1916
    |||::: ||| |||||
1016 GlnThrGlyValGlyGlyGlnIleGlyAlaGlyGlnProGlyTYrGlySe 1032
    |||::: ||| |||||
1917 GCGTCCGGCT.....GGATCTCAGCATTTGAGCTACGCGCT 1951
    |||::: ||| |||||
1032 rGlnProGlyIleGlyGlyGlnThrGlyAlaGlyGlnProGlyTYrGly 1049
    |||::: ||| |||||
1952 TGGCGCGTGGAGCGGTTTCACAGATTTGAGCTACGCGCTGCGCGGT 2001
    |||::: ||| |||||
1049 IagInProGlyPheGlyGlyGlnPro...GlyTYrGlyAsnGlnProGly 1064
    |||::: ||| |||||
2002 .....GGCTCTAGCCAGCATTTGAGCTACGCGCTGCGCGCT 2036
    |||::: ||| |||||
1065 ValGlyGlyGlnThrGlyAlaGlyGlnPro...GlyTYrGlySerGlnP 1080
    |||::: ||| |||||
2037 TGGCAGCGGTAGCCAGATTGAGC.....TACGCGCTGCGCTC 2074
    |||::: ||| |||||
1080 oGlyValGlyGlyGlnThrGlyAlaGlyGlnProGlyTYrGlyValIleP 1097
    |||::: ||| |||||
2075 CGGCT 2079
    |||::: ||| |||||
1097 roGly 1098
    |||::: ||| |||||

```

seq\_name: sp\_bacteria:Q91469

seq\_documentation\_block: PRELIMINARY: PRT: 1706 AA.

AC Q91469;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)  
 DE BIFUNCTIONAL HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (Ec 4.6.1.1).  
 GN CYAA.  
 OS Bordetella parapertussis.  
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
 OC Bordetella.  
 OX NCBI\_TaxID=519;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=63.2;  
 RA Boursaux-Ende C., Guiso N.;



"Sequencing of the Bordetella parapertussis adenylate cyclase-hemolysin-encoding gene."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249835; CAB76450.1; -  
 KW Lyase.  
 SO SEQUENCE 1706 AA; 177038 MW; 71750E2D0BFB64 CRC64;

## alignment\_scores:

Quality: 205.50 Length: 477  
 Ratio: 0.849 Gaps: 16  
 Percent Similarity: 50.734 Percent Identity: 22.432

## alignment block:

US-09-306-689-12 x Q9L469 ..

Align seg 1/1 to: Q9L469 from: 1 to: 1706

```

16 GATCGATCTCAGATCTGAGCTACGCGCTCGCGCTGCGAGCGGTTCTCA 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 ASPVALVALGlnHisGlyThrGlnAsnAspProPheProGlnAlaAs 310
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
66 AGATTGGAGCTACGCGCTCGCGCTGCGCTTACCCAGCATTTGAGCT 115
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
310 pGIuLysIlePheValValSerAlaThrGlyGlnMetLeuThr 327
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
116 AGGCG...CTGCGCCCTGCGAGCGGTACCA... 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
327 rGlglyGlnLeuLysGlnIleGlyGlnGlnArgGlyGlnGlyTyrVal 343
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
145 .....GATTGAGCTACGCGCTCGCGCTGCGATCTCAGCATTTG 185
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
344 PheTyrGlnAsnArgAlaTyrGlyVal.....AlaGlyLysSerLeu 358
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
186 GAGCTACGCGCTGCGCTGCGCTGCGAGCGTTCTCAAGATTGAGCTACGCGC 235
   ::::: ||||| ::::: |||
358 eAspAspGlyLeuGlnGlyAlaAlaProGlyVal..... 368
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 TCGCTCGCGCTGCTTACCCAGCATTTGAGCTACGCGCTCGCGCTGCGC 285
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 .....ProGlyGlyArgSerLys..... 374
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
286 AGCGGTACGCGCAT.....TGGAGCTACGCGCTGCGC 317
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
375 ...SerSerProAspValLeuGlnThrValProAlaSerProGlyLeu 390
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
318 TCGCGGTGATCTAGCTTCCCAAAACTGGGCAAAAAAATTATCTCT 367
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
390 gArg.....ProSerLeuGlyAlaValGlu..... 398
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
368 AATATCCCAAAATTACCAATATGATAGCAACAGTAATGCTTTACAG 417
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
399 .....ArgGlnAspSerGlyTyrAspSerLeuAspIleValGlySerArg 413
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
418 GATTTAGTCAAAAGCGGCCGAAGAGTTGGGGATTGAGGTACAAAGAGANA 467
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
414 SerPhe..... 415
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
468 ACGCATATATTTGCAACAGCTCAACAGATTAGCAGATTCAACCG 517
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 .....SerLeuGlyGlnValSerAspM 423
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
518 CTAATGGCTTAACAGCGCTGCGATTGTATCCGCTCCACAATTTGAT 567
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
423 et.....AlaAlaValGlnAlaValGlnLeuGln 432
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
568 AATATGCTACAGAAAAGTAAGAGCGCAAGCATTAAGTTAGTGGCGCAAG 617
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
433 MetThrArgGlnValLeuHisAlaGlyAlaArgGlnAspAspAlaLeu 449
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
618 CATTGTACAAATTCGAATTA.....GCCAAACTGTAATTATCTGCGCA 661
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
449 ogIlyValSerGlyAlaSerAlaHisTyrpIlyGlnArgAlaLeuGlnGly 466

```

```

662 TTCATCTATTTAGGCTCAGTATTGGCTGGAATGATGATTAGATGAGCC 711
   ||||| :::::
466 lAglnAlaValAlaAlaAla..... 472
   :::::
712 TTACAGATAACAGCAACCATGCTTGTCT..... 744
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
473 .....GlnArgLeuValHisAlaIleAlaLeuMetThrGlnPheG 486
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
745 .AAAGCTGCTTGAGCTTACAAATTCATTAATTTGAAAATTTGCTAAT 793
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
486 yArgAlaGlySerThrAsnThrProGlnGlnAlaAlaSerLeuSerAla 503
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
794 CAGTAAAAACACTTGACGAAATTTGGGACCAATTAATGCAATTTGGTTCA 843
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
503 lValAlaPheGlyLeuGlyGlnAlaSerSerAlaValAlaGlnThrValSer 519
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
844 AAACCTACAAATATCAAGCTTAGGACTTTAGGAGCAACACTCAAAA 893
   ||| ::::: |||
520 .....GlyPhePheArgGlySerSerArgTr 528
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
894 TATCGGTGACTTGAATTAAGCTGCGCTTGTAGATTTATCTCAGGCG 943
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
528 pAlaGlyGlyPheGlyValAlaGlyValAlaMetAlaLeuGlyGlyI 545
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
944 TATTATCGGGCCGACAGCTGCTTACTTTCAGATTAATAATGCTTCA 993
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
545 lAlaAlaAlaValAlaGlyAlaGlyMetSerLeuThrAsp...AspAlaPro 560
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
994 ACAGCTAAAAAGTGCGGCTTGTGAATGGCAACCAAGTTGTGG 1043
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
561 AlaGlyGlnLysAlaAlaValAlaGlyAlaGlnIleAlaLeuGlnLeuThr 577
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1044 TATATTTACCAAGCGCTTCTTCTTACATTTA.....GCCC 1081
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
577 yGlyThrValGlnLeuAlaSerSerIleAlaLeuAlaLeuAlaAla 594
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1082 AACGTTGTCAGCAGTTTATCTCACTGCGCGCTGCTGCTTAAAT 1131
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
594 rGlyValThrSerGlyLeuGlnValAlaGlyAlaSerAlaGlyAlaAla 610
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1132 GCTTCTACTGTTTCTTCTGCAATGACCATTAAGCATTTGCGGATTTGC 1181
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
611 AlaGlyAlaLeuAlaAlaAlaLeuSerProMetGlnIleTyrGlyLeuVal 627
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1182 CGATAATTTAATCATGCAAAAGTTTAGAGATTATGCCAGACCTTTA 1231
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
627 lGlnGlnSerHisTyrAlaAspGlnLeuAspLysLeuAlaGlnGlnSer 644
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1232 AAAAATTAGCTATGACGAGATTAATTTATAGCAAGATATCAGCGGGA 1281
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 eAlaTyrGlyTyrGlnGlyAspAlaLeuAlaGlnLeuTyrArgAsp 660
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1282 ACAGGACTATGATGATCGGTTACTGCAATTAATACCGATTGGCGGC 1331
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
661 LysThrAlaAlaGlnGlyAlaValAlaGlyValSerAlaValLeuSerThr 677
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1332 TATTGCTGCTGCTGCTGCTGCTGCTGCAACC 1362
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
677 rValGlyAlaAlaValSerIleAlaAlaAla 687
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq\_name: sp\_bacteria:Q9KX33

seq\_documentation\_block:

ID Q9KX33 PRELIMINARY; PRT; 3381 AA.  
 AC Q9KX33;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE SRPA.  
 OS Streptococcus cristatus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 Streptococcus.

OX NCBI\_TaxID=45634;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CC5A.  
 RA Correia F.F., Allen T.W., Dirienzo J.M.;  
 RT "High molecular weight serine-rich protein gene (srpA) from  
 RL Streptococcus crista."  
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U96166; AAF34780.1;  
 SO SEQUENCE 3381 AA; 321836 MW; C623EE363E74853P CRC64;

## alignment\_scores:

Quality: 199.00 Length: 714  
 Ratio: 0.463 Gaps: 23  
 Percent Similarity: 60.224 Percent Identity: 20.448

## alignment\_block:

US-09-306-689-12 x Q9KX33 ..

Align seg 1/1 to: Q9KX33 from: 1 to: 3381

```

26 AGCATTTGAGCTACGGCTCGCCCTGGCAGCGTTCTCAAGATTGGAGC 75
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2260 SerValSerAlaSerGlnSerAlaSerAlaSerAlaSerAlaSerAla 2276
   76 TACGGCTCGCTCGCGGTGCTAGCCAGCATTTGAGCTACGGCTGCG 125
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2276 nSerAlaSerIleSerValSerAlaSerAlaSerAlaSerIleSerAl 2293
   126 CCCGTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCGGTGCGGAT 175
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2293 erGlnSerAlaSerAlaSerIleSerAlaSerIleSerAlaSerIleSer 2309
   176 CTCAGCATTTGAGCTACGGCTCGCCCTGGCAGCGTTCTCAAGATTGG 225
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2310 AlaSerValSerAlaSerGlnSerAlaSerAlaSerValSerAlaSerG 2326
   226 AGTACGAGCGCTGCTCGCGGTGCTCTAGCCAGCATTTGAGCTACGG 275
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2326 nSerThiSer.....AlaSerValSerAlaSerGlns 2337
   276 GCGCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCGGTGCGG 325
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2337 erAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerIleSerVal 2353
   326 GATCTAGCTTCCCAAAACCTGGCGCAAAAAAATTATCTCTATATTGCC 375
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2354 Ser...AlaSerGlnSerAlaSerIleSerAlaSerGlnSerAlaSerAl 2369
   376 CAAATTTACCAATTTGATACCT...GAGCAAGTAAATGTTTACAGATT 421
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2369 aSerIleSerAlaSerIleSerAlaSerGlnSerAlaSer..... 2382
   422 TAGTCAAGCGCGCGAGAGATTGGGATTGAGTACAAAGAGAGAACGCG 471
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2383 ..ValSerAlaSerGlnSerAlaSerAlaSerValSerAlaSerIleSer 2398
   472 AATATATTTGACAGCTCAACCAAGCTTAGCCAGCATTTCAACCGCTAT 521
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2399 ThrSerAlaSerValSerAlaSerIleSerAlaSerAlaSerAlaSerAl 2415
   522 TGGCTTACCTGAGCGTGGCATTTGCTTATCCGTCACCAAAATTTGATAAT 571
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2415 nAsnGlnAsnSerAlaSerIleSerValSerAlaSerIleSerAlaSerI 2432
   572 TGCTACAGAAACTAAAGCAGCCAGCATTTAGTTTCCGGAAGACATT 621
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2432 IeSerAlaSerGlnSerAlaSerAlaSerAlaSerIleSerAlaSerIleSerAla 2448
   622 GTACAAATGCAAT.....AAAGCAAAACTGATTATTC 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2449 SerGlnSerAlaSerValSerAlaSerGlnSerAlaSerAlaSerValSe 2465

```

```

657 TGGCATTCAAATCTAATTTTAGCTCAGTATTTGCTGGAATGATTATGATC 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2465 rAlaSerGlnSerThrSerAlaSerValSerAlaSerGlnSerAlaSerAla 2482
   707 AGCGCTTACAGAACTAAAGCAGCCAGCATTCCTCTGCT..... 744
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2482 IaSerAlaSerAlaSerAlaSerGlnAsnGlnAsnSerAlaSerIleSerValSerAlaSer 2498
   745 ..AAAGCTGCGCTGGAGCTAACAAATTCATTAATGAAATATTTGCTAA 791
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2499 GlnSerAlaSerIleSerAlaSerGlnSerAlaSerAlaSerIleSerAl 2515
   792 TTCAGTAAAAACACTGACGAAATTTGGTGAGCAAAATTTGCAATTTGCT 840
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2515 aSerIleSerAlaSerGlnSerAlaSerValSerAlaSerIleSerGlnSerAla 2532
   841 .....TCAAAACTACAAATATCAAAAGCTTAGGACTTTAGCA 879
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2532 erAlaSerValSerAlaSerGlnSerIleSerThrSerAlaSerGlnSerAlaSer 2548
   880 GACAACTCAAAATATCGGTGAGCTTGATTAAGCTGCGCTTGGTTAGA 929
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2549 ThrSerMetSerAsn.....SerValSerAlaSerIleSerAlaSerG 2563
   930 TGTATCTCAAGGCTATTATCGGGCGCAACAGCTGCACTTTGACTTCAG 979
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2563 uSerValSerThrSerMetSerAsnSerValSerAlaSerIleSerAla 2580
   980 ATAAAAATGCTTCACA.....GCTAAAAAGTGGGTGCGGTTTGAA 1023
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2580 erValSerAlaSerThrSerMetSerAsnSerValSerAlaSerIleSer 2596
   1024 TTGGCAAAACCAAGTTGTTGTAATATTAACAAGCCGTTCTCTCTACAT 1073
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2597 AlaSerGlnSerAlaSerThrSerMetSerAsnSerValSerAlaSerI 2613
   1074 TTTAGCCCAACGTTGTGACAGAGTTTATCTCACTGAGGCTGTGCTG 1123
   |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2613 eSerAlaSer...ValSerAlaSerThrSerMetSerAsnSerValSerAla 2629
   1124 CTTTAATTTGCTTACTGTT.....TCTCTGCGATTAGCCCAATTGCA 1167
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2629 IaSerIleSerAlaSerValSerAlaSerIleSerThrSerMetSerAsnSerVal 2645
   1168 TTGCGCGTATGCGCAATTAATTAATCAATGCAAAAGTTTAGAGAGTTA 1217
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2646 SerAlaSerIleSerAlaSerValSerAlaSerIleSerThrSerMetSerAsnSe 2662
   1218 TGGCGAAGCTTTAAAAAATTAGCTATGACGAGATTAATTTATTTAGCAG 1267
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2662 rValSerAla.....SerIleSerAla 2670
   1268 AATATTCAGCGGGGACA.....GGGACTATTGATGATCGGTTACT 1308
   :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2670 erValSerAlaSerThrSerMetSerAsnSerValSerAlaSerIleSer 2686
   1309 GCAATTTATACCGCATTTGCGCGTATGCTGCTGCTGCTGCTGCTGCTGCT 1356
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2687 AlaSerGlnSerAlaSerThrSerMetSerAsnSerValSerAlaSerI 2703
   1357 .GCAGCCGATTTAATTTGAAAAAGTTAAACATTAATCTTGTC..... 1398
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2703 eSerAlaSerValSerAlaSerIleSerThrSerMetSerAsnSerValSerAla 2720
   1399 ..ATTCAGCAATAGCAAAAAAGAGAAAGTGCACATTTCAAAACTGGTTCCGA 1446
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2720 erIleSerAlaSerValSerAlaSerIleSerThrSerMetSerAsnSerValSer 2736
   1447 GAGCGTGAATTTGCTAAAGAGTGCCTAATTAATTAAGCAATTAAGATGA 1496
   :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
2737 AlaSerIleSerAlaSerGln.....SerAlaSerThrSerMetSerAs 2751

```

```

1497 GAAATTCGAAGAAATCATCGGTGCAAAATGGGAGCGGATCAACCTCAAGC 1546
2751 nserValserAlaserlIeserAlaserValserAlaserThser... 2766
1547 AAGTTGATGATCTTATWCGCAAAAGGTAAACGCAAAATATACCCAGATGAC 1596
2767 ..... MetSerAInSer 2770
1597 CATCAAAAGTTGTTGATPACTAGATGATTTGCTCAACATGACAAAATGT 1646
2771 ValserAlaserlIeserAlaserGlu.....serAlaserThserMe 2785
1647 GACAAACAGCTTAGTAACTAATCTCA..TCGTAGTGCATTTACCT 1693
2785 fserAnserValserAlaserlIeserAlaserGluSerlaserThrs 2802
1694 CGCTTAAGATTCGCAAAATGATTAAGTGGCTCCACGTCATCAATGTGAT 1743
2802 erMeSerAnserValserAlaserlIeserAlaserGluSerlaser 2818
1744 CAAGTTATCTTCCTTCATTTGCTAGGGGATCTCAGCATTTGGAGCTA 1793
2819 AlaserValserAlaserGluSerThserAlaserGluSer..AlaserT 2835
1794 CGGCTCGGCGCCCTGGCAGCGGTTCTCAAGATTGAGCTACGCGCTCGCTC 1843
2835 hSerMeSerAnserValserAlaserlIeserAlaserGluSerVal 2851
1844 CGGGTGGCTCA.....GCCAGCATTTGGAGCTACGCGCTG 1878
2852 SerThSerMeSerAnserValserAlaserlIeserAlaserValse 2868
1879 CGCCCTGGCAGCGGTAGCCAGATTGGAGCTACGCGCTCGCGGTG 1928
2868 rAlaserThSerMeSerAn.....serValS 2878
1929 ATCTAGCATTTGGACTACGCGCTGCGCCCTGGCAGCGGTTCTCAAGATT 1978
2878 erAlaserlIeserAlaserAla..... 2887
1979 GGAGCTACGCGCTGCGTCCGGGTGCTGTACCGACGATTGGAGCTACGGC 2028
2888 ...SerThSerleuSerAnserValserAlaserlIeserAlaserGl 2903
2029 CTGCGCCTGGCAGCGGTAGCCAGATTGGACACGCGCC 2068
2903 nserAlaserAlaserValserAlaserGluSerThser 2916

seq_name: sp.Bacteria:052781

seq_documentation_block:
ID 052781 PRELIMINARY; PRT; 1112 AA.
AC 052781:
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TReMBLrel. 06, Last annotation update)
DE SURFACE LAYER PROTEIN.
GN SAMP2.
OS Campylobacter fetus.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group.
OC Campylobacter.
OX NCBI_TaxID:196;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN:GIP 5396T;
RA Guesdon J L., Casadamek I., Chevrier D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF048699; AAC04580.1; -.
SQ SEQUENCE 1112 AA; 112503 MW; DA9D9C2F62602091 CRC64;

```

[illegible]

```

1259 sasnlle...glulysleuserleuthrasnserValSerasnArgt 275
1052 CCAAGACGGCTTCTTCTTACATTTTAAAGCCAGCTGTGCA..... 1092
275 hrheaspAlaLysGlyIleasplyleuGlnThrValAlaLeuSerGly 291
1093 ...GCAGGTATCTTCAACTGGCGCTGTGCTCTTAATGCTTAC 1139
292 GlulysGlyIleSerValThrAsn.....LeuAlaAsnII 303
1140 TGTTCCTCTTCCGATTTAGCCCATTTGCGGTAATGCCATAAT 1189
303 eValAspValGluLeuThrAsnLeuLys.....AlaAspLysP 316
1190 TTATTCATGCAAAAGTTTAGAGT...TATGCCAAGCGTTAAAAA 1236
316 heasn.....ValaspsrIleThraAlaAspLysValLeuAsp 328
1237 TTAGCTATGACGAGATATTTATTTAGCAGAAATATCAGCGGAGACAG 1286
329 GlySerAlaAspValGlnAsnLeuLysVal.....AsnGlyValGI 342
1287 GACTATTGATGATCGGTACT.....GCAATTATATCCGCAT 1324
342 yAlaLysGlyAlaSerValAlaIleThrAlaAspLysIleGluThrLeuA 359
1325 TGCGCGCATTTGCTGTGCTGTCTGCTGCTGCGCATTTAAACATTT 1374
359 snLeuAsnThrThrGlySerGlnSerPheValSerAlaAspValAlaSer 375
1375 GAAAAGCTTAAACATATCTTGTATCAGCAATGCAAAAAAGAGAAAGT 1424
376 IleSerValLysGlyAsnAlaAsnLeuSerLeuAlaThrGlyAlaLysTh 392
1425 GACCATTCAAAACTGTTCGAGAGGCTGATTTGCTAAAGAACTGCCTA 1474
392 rThrThrLeuAsp.....AlaSerPheGlyGlyAlaLeuAspA 406
1475 ATTATAAAGCACTAAAGATGAGAAATCGAAGAAATCATCGTCAAAAT 1524
406 laAspLeuSerThr...SerAlaSerValThrSerIleLysGlyLysn 421
1525 GGC...GAGCGGATCACCTCAAGCAAGTT..... 1551
422 GlyAsnAspLysIleThrIleLysAspValAlaValAsnValAlaIleAs 438
1552 .....GATGATCTTATCGCAAAAGGTACGCGCAAAATTA 1585
438 pGlyGlyAlaGlyAsnAspGluLeuValIleLysGlySer.....T 452
1586 CCCAAGATGAGCTATCAAAAGTTGTGATACACTAGAA..... 1623
452 hrAlaAspThrLeuGlnProThrLeuThrAsnIleGlyLysValThrVal 468
1624 .....TTGCTCAACATACCAAAATGT 1646
469 AspGlyAsnThrLysAspLeuThrLeuSerLeuLysLysAlaGlnSerVa 485
1647 GACCAACAGCTTACATAAGTAAATCATCTGTAGTAGCAATTTGCTAGGAGT 1778
485 lThrGluSerSerPheLysAsnIleAlaLysThrValThrGluSerAsnG 502
1697 CTAT...GATTGAGAAATGATTATG.....GCTCCA 1728
502 LysAsnValGluThrValAsnIleLeuAlaAsnAsnAlaThrAspLysAla 518
1729 ACTTCATGTTGATCAAAAGTTTATCTCTCTCAATTTGCTAGGAGT 1778
519 ValThrIleAsnAspGluSerLeuLysThrIleAsnPheSerAspValAs 535
1779 TCACAGATTGAGCTACGCGCTGCGCCCTGCGAGGGGTTCTCAAGATTGCA 1828

```

```

535 pasPlys.GlyAlaSerValArgAlaLysGly.....LysIleVal 548
1829 GCTACGGCGCTGCTCGCGGAGGCTCTAGCCAGCTGAGCTAGCGGCTG 1878
549 AlaAspLysAlaThrGluLeuThrIleAsnSerAsnLysValThrAlaAl 565
1879 CGCCCTGCGACCGGTAGCCAGATTGAGCTACGCGCTGCGTCCGGCTG 1928
565 aAlaAspAlaValValGlnAlaAlaAsnAlaThr.....LysIleA 579
1929 ATCTC.....AGCATTTGAGCTACGCGCTCGCCCTGGCA 1963
579 spIleAsnAlaAlaLysAspThrValGlyLeuThr.....LeuGly 592
1964 GCGGCTTCAAGATT...GGAGCTACGCGCTGCGCTGCGGAGCTCTAGC 2010
593 GlyValAlaLysLeuThrAspLeuThrValAsnAsnLysGlyAlaPheAl 609
2011 CAGCATTTGAGCTACGCGCTGCGCTGCGGAGCTGAGCCAGATT 2056
609 aleuThrGlyAlaAsnAlaThrAspLeuAspSerValLysAsnLeu 624

```

seq\_name: sp\_r0dent:Q61869

```

seq_documentation_block:
ID Q61869 PRELIMINARY; PRT; 707 AA.
AC Q61869;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE KERATIN COMPLEX 2, BASIC, PROTEIN 2 (KERATIN 2 EPIDERMIS).
GN KRT2-17 OR KRT2-2 OR MK2E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE=94149286; PubMed=7508961;
RT "The large type II 70-kDa keratin of mouse epidermis is the ortholog
RT of human keratin K2e."
RL J. Invest. Dermatol. 102:165-170(1994).
CC -1- SIMILARITY: TO ALL OTHER INTERMEDIATE FILAMENT PROTEINS.
DR EMBL: X74784; CAA52788.1; -.
DR MGD: MGI:96699; Krt2-17.
DR INTERPRO: IPR001664; -.
DR INTERPRO: IPR003054; -.
DR PFM: PF00038; filament; 1.
DR PRINTS: PR01276; TYPE2KERATIN.
DR PROSITE: PS00226; IF. 1.
KW keratin; intermediate filament; Coiled coil; Heptad repeat pattern.
SQ SEQUENCE 707 AA; 70977 MW; D58FA5E6E30BB86 CRC64;

```

alignment\_scores:

```

Quality: 187.50 Length: 750
Ratio: 0.593 Gaps: 38
Percent Similarity: 42.133 Percent Identity: 21.867

```

alignment\_block:

us-09-306-689-12 x Q61869 ..

Align seg 1/1 to: Q61869 from: 1 to: 707

```

91 GGTGCTCTACAGCTAGCTAGCTAGCGCTGCGCTGCGAGCGGTAG 140
|||||
86 GlyLysSerSerPheGlyLysSerSerGlyPheGlyGlyAlaGlyPh 102
141 CCAAGATTGAGCTACGCGCTGCGCTGCGGATCTCAGCATTTGG... 186
102 e.....GlyGlyGlyGlnGlyPheGlyG 110

```

```

187 .. AGCTACGGCTCGCCCTGGCAGC.....GGTCTCAAGATTGG 225
    ||| ||||| ||||| ||||| ||||| |||||
110 LysSerGlyPheGlyGlyGlySerGlyPheGlyGlyGlnGlyPhe 126
    ::| ||| ||| ||||| ||||| |||||
226 AGCTACGGCTCGCTCCGGGTGGCTTACCCAGCATTTGAGTACGGCT 275
    ::| ||| ||| ||||| ||||| |||||
127 GlyGlyGlySerArgPheGlyGlySerGlyPheGlyGlyGlyPhe 143
    : ||||| ||||| ||||| |||||
276 GGGCCCTGGCAGC.....GTAAGCCAGATTGGAGCTACGGCTCGCTC 319
    : ||||| ||||| ||||| |||||
143 eGlyGlyGlySerPheGlyGlyArg.....PheGlyGlyGlyPhe 157
    : ||||| ||||| ||||| |||||
320 CGGGTGA.....TCTAGCTCCCAAAACCTGGGGCAAAAAA 357
    ||||| ||||| ||||| |||||
157 roGlyGlyPheGlyGlyProGlyGlyPheProGlyGlyGlyIleHisGly 173
    ||||| ||||| ||||| |||||
358 AAT.....ATCCTATATATCCCAAAATTTACCAATATGATAC 395
    ::| ||| ||| ||| ||| |||
174 ValSerValAsnGlnSerIleuGlnProLeuAspValLysValAspPr 190
    ::| ||| ||| ||| ||| |||
396 TGAACAAGTAATGGTTTACAGGATTACTCAAGCGCCGAAGAGTTGG 445
    ||| ||| ||| ||||| |||||
190 oGluIleGlnAsn.....ValLysSerGlnGluArgGluG 202
    ||||| ||||| ||||| |||||
446 GGATTGGGTACAAAGA..... 462
    ||||| ||||| ||||| |||||
202 IuIleLysThrLeuAsnAsnLysPheAlaSerPheIleAspLysValArg 218
    ||||| ||||| ||||| |||||
463 .....GAAGAACGCAATATATTTGCACAGCTCAAAACAGTTTA..... 501
    ||||| ||||| ||||| |||||
219 PheLeuGlnGlnGlnAsnGlnValLeuArgThrLysTrpGluLeuLeuL 235
    ||||| ||||| ||||| |||||
502 .....GGCAGCATTCAAA 514
    ||||| ||||| ||||| |||||
235 nGlnLeuAlaCysArgGlnProHisIleLysProAspProIlePheGlnA 252
    ||||| ||||| ||||| |||||
515 CGGTAATGGCTTAACCTGAGCGGTGATGTGTATCCGCTCCACAAT 564
    ::| ||||| ||||| |||||
252 IeTyrlleGlyMetLeuLysLys.....GlnVal 261
    ||||| ||||| ||||| |||||
565 GATTAATTTG...CTACAGAAACTAAAGCAGCAGCATTTAGGTTCTGC 611
    ||||| ||||| ||||| |||||
262 AspArgLeuSerAlaGlnArgThrSerGlnGluSerGluLeuAsnAsnMe 278
    ||||| ||||| ||||| |||||
612 CGAAAGCATTTGACAAAT.....GCAAAATA 637
    ||||| ||||| ||||| |||||
278 tGlnAspLeuValGluAspPheLysLysLysTrpGluAspGluIleAsnL 295
    ||||| ||||| ||||| |||||
638 AAGCCAAACTGATTTATCTGGCATTTCAATCTATTTAGGCTCAGATTTG 687
    ||||| ||||| ||||| |||||
295 yAspThrSerAlaGluAsnAspPheValThrIle..... 306
    ||||| ||||| ||||| |||||
688 GCTGAATGATTTAGATGAGCGCTTACGAATTAACAGCAACCAACATCG 737
    ||||| ||||| ||||| |||||
307 ..LysLysAspValAspSerCysTrpMetAsp..... 316
    ||||| ||||| ||||| |||||
738 TCTTCCTAAAGCTGGCTGGAGCTAAACAATTCATTATTTGAAATATTG 787
    ||||| ||||| ||||| |||||
317 .....LysThrGluLeuGlnAlaArgLeuAspIleLeuAlaGlnGluV 331
    ||||| ||||| ||||| |||||
788 GTAATTCACTAAACACTTGACGAATTTGGTGAGCAATTTAGTCAATTT 837
    ||||| ||||| ||||| |||||
331 AlAspPheLeuArgThrLeu.....TyrAspAlaGluLeuSerGlnLeu 345
    ||||| ||||| ||||| |||||
838 GGTTCAAAATCTACAAAATATCAAGGCTTAGGACTTTAGAGACAACACT 887
    ::| ||| ||| ||||| ||||| |||||
346 GlnGlnAspValThrAspThrAsnValIleLeuSerMet...AspAsnTh 361
    ::| ||| ||| ||||| ||||| |||||
888 CAAAATATCGGTGACCTTGATTAAGCTGGCTTGGTTAGATGTTATCT 937
    ::| ||| ||| ||||| ||||| |||||
361 rArgAsn.....LeuAspLeuAspSerIleI 370

```

```

938 CAGGGCTATTATCGGGCGCAACAGTGCATTTCTACGTACGATAAAT 987
    ::|
370 IeAla..... 371
988 GCTTCAACAGCTAAAAAGTGGTGGGCTTTGAAATTTGGCAACCAAGT 1037
    ||||| ||||| |||||
372 .....GluValGlnAsnGln.. 376
1038 TGTGGTATATTACCAAGCCGTTTCTTTACATTTTAGCCCAACGAG 1087
    ||||| ||||| ||||| |||||
376 ..... 376
1088 TTGCACAGGTTTATCTCACTGGGCTGTGGCTGCTTTAATTCCTCT 1137
    ||||| ||||| ||||| |||||
376 ..... 376
1138 ACTGTTTCTTTCGATTAAGCCATTAGCATTTGCCGTAATGCCGATA 1187
    ::| ||||| ||||| |||||
377 .....TyrGluMetIleAlaHisLys 383
1188 ATTTAATCATGCAAAAGTTTAGAGATTATGCCGAACGCTTTAAAAAT 1237
    : ||||| ||||| ||||| |||||
383 s...SerLysAlaGluSerGluGluLeuTyrlHisSerLysTrpGluGluL 399
    : ||||| ||||| ||||| |||||
1238 TAGGCTAT.....GACGAGATTAATTTATACAGAAATATCAG 1275
    ||||| ||||| ||||| |||||
399 eGlnValThrAlaValLysHisGlyAspSerLeu..... 410
1276 CGGGGAACAGGCAATTTGATGATCGATCGGTTACTGCAATTAATCCGAT 1325
    ||||| ||||| ||||| |||||
411 .....LysGluIleLysMetGluIleSerGluLeuAsnArgThrI 424
1326 GCGCGTATTCCTGGTGGTGTGTCTGCTGCGACCGCATTTAACATTTG 1375
    : ||||| ||||| ||||| |||||
424 eGlnArgLeuGlnGlyGluIleSer..... 432
1376 AAAAGTTAAACATATCTTTCATCAGCAATACCAAAAGCAAGAG 1425
    ::| ||||| ||||| |||||
433 .....HisValLysLysGlnCysIle 439
1426 ACCATTCAAAAGCTGTTCCGAGAGCGTGATTTTGTAAAGAGTGGCTAA 1475
    ::| ||||| ||||| |||||
440 GlyValGlnAspSerIleAlaAspAlaGlu..... 449
1476 TTATTAAGCACTAAAGATGAGAAAATCGAAGAAATCATCGTCAAAATG 1525
    ||||| ||||| ||||| |||||
450 .....GlnHisG 452
1526 GCGAGCGGATCACCTCAAGCAAGTTGATGATCTTTTCGCAAAAGCTAAC 1575
    ||||| ||||| ||||| |||||
452 LysGluHis.....AlaIleLysAspAlaArg 460
1576 GCGAAAATTAATTC.....CAAGATGAGCT 1598
    ||||| ||||| ||||| |||||
461 GlyLysLeuThrAspLeuGluGlnAlaLeuGlnGlnCysArgGluAspLe 477
1599 ATCAAAAGTTGTGATTAATGATGAAATTTGCTCAACATACCAAAATGTGA 1648
    : ||||| ||||| ||||| |||||
477 uAlaArgLeuLeuArgAspTrpGlnGluLeuMetAsnThrLys..... 491
1649 CAACAGCTTAGAT.....AAGTAAATTCATCT 1677
    ||||| ||||| ||||| |||||
492 ..LeuSerLeuAspValGluIleAlaThrTyrArgLysLeuLeuGluGly 507
1678 GTAAGTCATTTTACCTGCTTAATGATTTGAGAAAATGATTAATGGCTGC 1727
    ::| ||||| ||||| ||||| |||||
508 GlnGluCysArgMetSerLysAspHisSerAspAsnValSerValSerI 524
1728 AACTTCATGTTGGATCAAGT.....TTATCTCTCTTCATTAATTTGCTA 1771
    ||||| ||||| ||||| ||||| |||||
524 eThrSerSerThrIleSerSerMetAlaSerLysThrGlyPheGlyLys 541
1772 GGGGATCTCAG.....CATTTGAGCTACGGCTGGCCCTGGC 1809

```



```

579 aileasplalaileuasnllysAlaSerSerAlaasnThrAspArgile.... 594
941 GGCTATTATCGGGCCACAGCTGCACATTGTACTTGCAGAT.....AAA 984
595 .....AlaThrAlaIleuLeuGlyIleAlaGluAsnLys 606
985 AATGCTCAACAGCTAAAAAGTGGGTGGGCTTTGAAATTGCAACCA 1034
607 AspAlaGlnIleAlaLys.....AlaGlnAlaAsnGln 617
1035 AGTTGTGGTAATATTATACCAAGCGCTTCTTACATTATAGCCCAAC 1084
617 uasnllysAspGlyIleAlaLysAsnGlnAlaAspIleGlnLeuHisAspL 634
1085 GTGTTGCAGCAGCTTTATCTTCAACTGGCGCTGGCTGCTTAATGCT 1134
634 yslAspIleThrAsnLeu.....GlyIleLeuHis 643
1135 TGTACTGTTTCTGGCATTAGCCCATTAATGATTTGGCGTATTGGCA 1184
644 SerMetValAlaArgAlaValGlyAsnAsnThrGlnGlyValAlaThrAs 660
1185 TAAATTTATCATGCAGAAAGTTTAGAGATTATGCCGAACGCTTAAAA 1234
660 nlysAlaAspIleAlaLysAsnGlnAlaAspIleAlaAsnAsnIleLysA 677
1235 AATTAGCTATGACGAGATTAATTATTAGCAGAAATATCAGCGGGACAA 1284
677 smile...TyrGluLeuAlaGlnGlnGlnAspGlnHisSerSerAspIle 692
1285 GGGACTATGATGATCGCTGCTACTGCAATTAATACCGCTTGGCGCTAT 1334
693 LysThrLeu...AlaLysValSerAlaAlaAsnThrAspArgIleAla... 707
1335 TGCTGTGGTGTGTCTGCTGTCGACCGCATTTAATCTGAAAAAGTTA 1384
708 .....LysAsnLysAlaGlnAlaAspAlaSerPheGlnThrLeuT 721
1385 AACATATCTTGTCAATCAGCAATAGCAAAAAAGCAAAAGTCAACATTCA 1434
721 hLysAlaAsnGlnAsnThrLeuIleGlnGlnGlnAlaLeuValGlnGln 737
1435 AACGTGTTCCGAGAGCGTATTGCTTAAGAAGAGCCCAATATTATAAC 1484
738 Asn.....LysAlaIleAsnGlnGlnLeuGlnGlyPheAlaAla 750
1485 AACATAAGATGAGAAATGCAAGAAATCATCGTCAAAATGCGCAGCGGA 1534
750 hAlaAlaAspValGlnAspLysGlnIleLeu...GlnAsnGlnAlaAspI 766
1535 TCACCTCAAAAGCAAGTGTGATCTTTATGCAAAAAGTAAAGCGCAAAATT 1584
766 leThrThrAsnLysThr.....AlaIle 773
1585 ACCCAAGATGAGCTATCAAAAGTGTGATTAATCACTATGCTGCTCAAAA 1634
774 GlnGlnAsnIleAsnArgThrValAlaAsnGlnGlyPheGlnIleGluLys 790
1635 TAGCAAAAATGTGACAAACAGCTTAGATTAAGTAAATCTCATCTGTAAAGT 1684
790 nLysAlaGlyIleAlaThrAsnLysGlnIleLeu..... 802
1685 CATTTACTCTGTAATGATTCGAGAAATGTATTAGTGGCTCCAACTTCA 1734
803 .....LeuGlnAsnAspArgLysAsnArgIleAsnGlnLysThrAsnAsn 816
1735 AATGTGATCAAGATTATCTCTCTCAATTTGCTAGGGAATTCACAGA 1784
817 hLysGlnAspGlnLysIleAspGlnLeuGlyTyrAlaLeuLysGlnIle 833
1785 TTGGAGATCGCGCTGCGCCGTCAGCGTCTCAAGATTGAGACTGAC 1834
833 yGlnHisPheAsnAsnArgIleSerAlaValGlnArgGlnThrAlaGlyG 850

```

```

1835 GCCTG.....CGTCCGGGTGGCTCT 1854
850 lYlLeAlaAsnAlaIleAlaIleAlaThrLeuProSerProSerArgAla 866
1855 AGCCAGCAT.....TGAGCTACGCGCTGCGCCCTGACGACGCGTAG 1895
867 GLyGlnHisHisValLeuPheGlySerGlyTyrHisAsnGlyAlaAla 883
1896 CCAAGAT...TGAGCTACGCGCTGCGCCGCGGTGATGATTCAGCATTTGA 1942
883 aValSerLeuGlyAlaAlaGlyLeuSerAspThrGlyLysSerThrTyrL 900
1943 GCTACGCGCTGCGCCCTGCGCAGCGGTCTCAAGATTGGAGC.....TAC 1986
900 yslleGlyLeu.....SerThrAspAlaGly 909
1987 GGCCTGCTCGCGGT...GCCTTACCGCAGCATTTG 2019
910 GlyLeuSerGlyGlyValGlyGlySerTyrArgTyr 921

seq_name: sp_invertebrate:Q9YTR6
seq_documentation_block:
ID Q9YTR6 PRELIMINARY; PRT; 1039 AA.
AC Q9YTR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CG5700 PROTEIN.
GN CG5700.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottler P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris J.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaiswal M., Kalush C., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paolel J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

```

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003543; AAF4980.1; .  
 DR FLYBASE: FBgn0036231; CG5700.  
 SO SEQUENCE 1039 AA; 102635 MW; 6A710871139BD01F CRC64;

alignment\_scores:  
 Quality: 185.50 Length: 629  
 Ratio: 0.602 Gaps: 31  
 Percent Similarity: 48.967 Percent Identity: 22.258

alignment\_block:  
 US-09-306-689-12 x Q9VTR6 ..

Align seg 1/1 to: Q9VTR6 from: 1 to: 1039

```

37  TACGGCGCTGCGCCCTGACAGCGGTCTCAAGATTGGAGC..... 75
323 TyrGlySerGlnProGlyIleGlyGlnThrGlyAlaGlyGlnProGly 339
76  .TACGGCGCTGCGCTGCGGTGCTTACCCAGCATTTGAGCTACGCGCTGC 124
339 yTyrglyAlaGlnProGlyPheGlyGln...LeuGlyTyrclyAsnG 355
125 GCGCTGACAGCGGTACCAAGATTGGAGC.....TACGGCGCTG 162
355 InProGlyValGlyGlyGlnThrGlyAlaGlyGlnProGlyTyrglySer 371
163 CGTCCGGGT.....GGATCTCAGCATTTGAGCTACGCGCT 197
372 GlnProGlyAlaGlyGlnThrGlyAlaGlyGlnProGlyTyrglyVa 388
198 GCGCGCTGACAGCGGTCTCAAGATTGG..... 225
388 ILeProGlyPheGlyGlyGlnProGlyIleGlyGlnThrAlaAlaG 405
226 .....ACCTACGCGCTGCGTCCGGGT...GGCTTACCCAGCATTTG 264
405 lTyrsProGlyTyrglyGlyGlnProGlyIleGlyGlySerPro..... 419
265 AGCTACGCGCTGCGCGCTGACAGCGGTACCAAGATTGG..... 303
420 ValTyrclyThrGlnGlnGlyThrGlyGlyGlnSerGlyIleSerGly 436
304 .....AGCTACGCGCTGCGTCCGGGTGATCT..... 330
436 yGlnProGlyTyrclyThrGlnProGlyGlnThrGlyAlaGlyGlnPro 453
331 .....ACCTTCCCAAAAACCTGGGCAAAAAAATTATC..... 363
453 lTyrclySerLeuProGlyThrGlyGlnAlaThrAlaGlyGlnPro 469
364 CTCATATATTCCTCCCAAAATTACAA..... 387
470 GLyTyrclyProGlySerGlnProGlyIleGlyGlyGlnThrValGly 486
388 ..... 391
486 yHisGlyGlyTyrclySerGlnProGlyIleGlyGlyAlaProValTyrg 503
392 ATACTGACAAGTAATGCTTTACAGATTATTCAAAGCGCGGAA... 438
503 lYthrGlnProGlyGlyGlyGlnThrGlyValIleGlyGlyGlnPro 519
439 ...GAGTTGGGATTGAGTACAAGAAAGCAATATATTTATTCGAC 485
520 GLyGlnIleGly.....AspArgValGlyGlnProGlyTyrglyTh 533
486 AGCTCAAAACGATTTAGCAGC.....ATTCAAA 514

```

```

533 rGlnThrGlyGlnIleGlyAlaProGlyArgTyrrhrAspGlySerGlnr 550
515 CCCTATTGGCTTAACAGCAGCGGCTGCTTATCCGCT.....CCA 558
550 hrValProGlyAlaValGlyThrGlyGlyValAlaAlaGlyThrSer 566
559 CAATTGATTAATTTGCTACGAAACTAAAGCA.....GCCCA 596
567 GLyAlaAspAlaPheSerGlnAlaGlySerGlnIleGlyAspGly 583
597 AGCATTAAGTTCGCCGAAGCATTTGCAAAATGCAAAATGCAAAATG 646
583 nAlaSerAlaSerAlaGlnGly...LysLysAsnGlyGlyThrAlaLys 599
647 CTGATTAATCTGCATTCATTCATTTAGCTCAGATTTAGCTGCAATG 696
599 hrGlnValSerGlyThrTyrrSerSerGlyGlyThrPheSerAla..... 613
697 GATTTAAGTAGGCTTACAGAAATACAGCAACCAACATGCTTGTCTAA 746
614 .....SerAlaMetThrSerAlaAspArgAlaAlaSerAlaGly 627
747 AGCTGCTTGACCTTAACAATTCATTAATGAAATATTCATTAATTCAG 796
627 nValThrGlyAsnAlaAspGlyAlaValSerGlnSerGlnGlySerGly 644
797 TAAAAACCTTGACGAATTTGGTGACCAATTTGTCATTTGCTTCAAAA 846
644 lYProAlaGlnSerGlnAlaGlnValAlaGlnAlaLysAspGlyGlyThr 660
847 CTACAAAATATCAAAAGCTTAGGCACTTTAGGACAACTCAAAATAT 896
661 LysAlaSerSerGlnSerGlyIleIleGlnGlnSerGlnSerGlnVa 677
897 CGGTGACCTTAATTAAGCTGCGCTTGAATGATTAATTCAGGCTAT 946
677 lHisAlaAsnAspLysGlyGly..... 684
947 TATCGGCGCAACAGCTGCACTTGTACTGCAATTAATAATGCTTCA... 993
685 .....LeuAlaAspAlaGlnSerSerGly 692
994 .....ACAGCTAAAAAAGTGGTGGCGGTTT.....GAAAT 1025
693 ProGlyGlnThrSerSerGlnAlaGlnIleGlyPheArgProGlyGln 709
1026 GGCAAACCAAGTGTGTGTAATATTACCAAGCCGTTCTTCTTACATTT 1075
709 uAlaAsnProIleAlaAlaAsnGlyGlyGlyGlnAlaSerSer..... 723
1076 TAGCCCAACGTTGTCAGAGGTTTATCTCACTGAGCGCTGGTGTCT 1125
724 .....SerSerGlyThrHisSerSer...GlnSerSerSer 734
1126 TTAATGCTTCTACTGTTCTCTGTCGATTAGCCCATTAACATTTGCCG 1175
735 GlnIleHisGlyThrSerSerPheGlyValSer.....TyrrHisGly 748
1176 TATTGCGCATTAATTTAATCATGCAAAA...AGTTTGAGAGATTATCCG 1222
748 yAlaAlaGlnSerAlaSerGlyThrLysGlnGlnValAlaThrTyrrArg 765
1223 AACGCTTAATAAATTTAGGCTATGACGAGATTAATTAATAGCAATAT 1272
765 lValAlaAsnArgGlnLeu.....PheAsnThrIleSerGlnPhe 777
1273 CAGCGGGGAACAGGACTATTGATGATCGATCGATTACTCAATTAATACGC 1322
778 GLyAsnAsnAlaAsnAlaValAlaThrAspArgAlaAspAlaValTyrrSer 794
1323 ATTGCGCGCTAATTCGTGCTGCTGCTGCTGCTGACCGCATTTAATCAT 1372

```



1991 AGGCCCTACTCCATCAATCTTGAGAACCGCGTCC.....AGGCGGAG 1951  
 108 rolaaThrrAlaalaThrProalaThrAlaAlaPheAsnPheAlaAlaThr 124  
 1950 GCCGTAGCTCCAAATCTGAGATCCACCGGAGCGAGCGGTAGCTCAAT 1901  
 125 AlaAlaThrProAlaThrAlaAlaThrAlaAlaLeuIlePheAlaAlaTh 141  
 1900 CTTGGCTACCGCTGCCGAGGGCGCAGCGGTAGCTCAATGCTGGCTAG 1851  
 141 rAlaAlaThrAlaAlaThrProAlaThrProAlaLeuAsnPheAlaAla 158  
 1850 CCACCCGGAGCGAGGCGGTAGCTCCAACTTGAGAACCGCGTCCAGGGC 1801  
 158 hrAlaAlaThrAlaAlaThrAlaAlaThrProAlaThrAlaAlaLeuAsn 174  
 1800 CAGCGCGGTAGCTCCAAATGCTGAGATCCCTCCAGCAATTTGAGAGAGAT 1751  
 175 pheAlaThrAlaAlaAlaThrAlaAlaThrProAla.....185  
 1750 AACCTTGATCCAACTTGAAGTTGGAGCGCACTAATATTCGAAATCA 1701  
 186 .....ThrProAlaLeuAsnPheAlaAlaThrAla 195  
 1700 TTAGACGAGGTAAATGCACCTTACAGATGAGATTAACCTTAACTGCTT 1651  
 196 AlaThrProAlaThrAlaAlaThrProAlaLeuAsnPheAla..... 209  
 1650 TGTCAACATTTTGTGATGTTGAGCAATTCATAGTATTCACCAACTTGTG 1601  
 210 .....AlaThrAlaAlaThr 214  
 1600 ATAGCTCATCTTGGGTAAATTTTGCGCTTACCTTTGCGATTAAGATCA 1551  
 214 hrAlaAlaThrProAlaThrAlaAlaLeuAsnPheAlaAlaThrAlaAla 230  
 1550 ACTTGCTTTGAGGTAGTCGCCGCTCGCCAAATTTGACCGATGATTTCTGAT 1501  
 231 ThrAla..... 232  
 1500 TTTTCATCTTAGTGTGCTTTATATATAGCACTTTCACCAAAATCAG 1451  
 233 .....AlaThrProAla 236  
 1450 CCTCTCGGAAACAGTTTGAATGACACTTTCCTTTTGTCTATGCTGTG 1401  
 236 IaThr.....AlaAlaPheAsnPhe..... 242  
 1400 ATGACAAAGATTAATGTTTAACTTTCAAAATGTTAAATCGGCTGAGCAGC 1351  
 243 .....AlaAlaThrAla 246  
 1350 AGACACACACACAGAAATAGGGCCAAATGGCGTATATATTCAGATACCG 1301  
 246 AlaIaThrProAlaThrAlaAlaThrProAlaLeuIlePheAlaAlaThr 263  
 1300 ATGCAATCAATAGTCCCTGCTCCCGCGCTGATATTCGCAAAATAATTAATC 1251  
 263 IaAlaThr.....ProAlaThrAlaAlaThr 271  
 1250 CCGTCAATAGCCTAATTTTAAAGCGTTCGGCATTAACCTCTAAACTTTT 1201  
 272 ProAla.....LeuAsnPheAlaThrAlaAlaThrAlaAlaThrProAlaTh 287  
 1200 TGCATGATTAATTTATTCGCAATACCGGCAAAATGCTAATGGCTAATCG 1151  
 287 rProAlaLeuAsnPheAlaAlaThrAlaAlaThrAla.....AlaThrAla 303  
 1150 CAAGAGAAACAGTAAAGCAATTTAAAGCAGACAGCGCCAGTGAAGAT 1101  
 303 IaThrProAlaThrAlaAlaLeuAsnPheAlaAlaThrAlaAlaThrAlaAla 319

```

1100 AACCTGCTGCAACGCTTGGGCTAAATGTAGAGAAAGCGCTTGGT 1051
    |||||
320 ThrProAlaThr..... 323
1050 AATATTACCAACACTGGTTGGCAATTCAAAACCGCACCTTTT 1001
    |||||
324 .....ProAlaLeuAsnPheAlaAlaThrProAlaThr 338
1000 TAGCTGTGAAGCTTTTATCTGCAATGCAAGTGA..... 963
    |||||
338 laAlaThrProAlaLeuMetPheAlaAlaThrAlaAlaThr 354
962 ...GCTGTGGCCCGATATAGCCCTGAGATACATCTAAACCAAGGCC 916
    |||||
355 ProAlaThrAlaAlaPheAsnPheAlaAlaThrProAlaThr 371
915 AGCTTTTCAGAGTCCACCGATATT.....TTGAGTTTGTCTC 878
    |||||
371 rAlaAlaThrProAlaLeuIlePheAlaAlaThrAlaAlaThrAla 388
877 CTAAAGTCCCAAGCCTTGATATTGTGAGTTGAACCA..... 837
    |||||
388 hrAlaAlaThrAlaAlaLeuIlePheAlaThrAlaAlaThrAla 404
836 .....AATTGACTAATTGCTCACCAAAATGCTCAG 805
    |||||
405 ProAlaThrAlaAlaPheAsnPheAlaAlaThrAlaAlaThrProAla 421
804 TGTATTACGCAATTAGCAATATTTCATTAATGAATTTGTGCTCCA 755
    |||||
421 rAlaAlaThrProAlaLeuIlePheAlaAlaThrAlaAlaThrAla 438
754 AGCAGCTTGAAGAGCATGTGTTGCTATTCTGTGAAGCCCTCA 705
    |||||
438 hrProAlaThrAla.....AlaLeuAsnPheAlaThrAlaAla 450
704 TCTAATCCATTCAGCCCAATCTAGCCCTAAATGATGAATGCCAGA 655
    |||||
451 ThrAlaAlaThrProAlaThr..... 457
654 TAATACAGTTTGGCTTATTGTCATTGTCACATGCTTTCGCAGAC 605
    |||||
458 .....AlaAlaLeuAsnPheAlaThrAlaAlaThrAlaAla 470
604 CTATGCTGGCGCTTACTTGTGTCAGCAATTAATCAATTTGCGA 555
    |||||
470 hrProAlaThrProAlaLeuAsnPheAlaAlaThrAlaAla..... 483
554 GCGGATACACAAATGCCAGCTCAGTTAAGCAATAGCGTTGCAATCGT 505
    |||||
484 .....Th 484
504 GCGTAACCTGTTGAGCTGTGCAATATTATTCGCTTCTCTTGTGA 455
    |||||
484 rProAlaThrAlaAlaThrProAlaLeuMetPheAlaAlaThrAlaAla 501
454 CCGCAATGCCCAACCTTGGCGCTTGAATAATCCGTGAACCAATTA 405
    |||||
501 hrAlaAlaAlaThrProAlaThrAlaAlaPheAsnPheAlaAla..Thr... 515
404 CCGTGTTCAGTATCATATTGGTAATTTGGGGAATATAGAGATAATTTT 355
    |||||
515 ..... 515
354 TTTTGGCCCGACTTTTGGGAAGTAGATCCACCGGACGACGCCGTAGC 305
    |||||
516 .....AlaAlaThrProAlaThrAlaAlaAla 524
304 TCCCATCTG.....GCTACGGTGCAGG...GCGGAGCCGCTAG 267
    |||||
524 rProAlaLeuIlePheAlaAlaThrAlaAlaAlaThrProAlaThrAlaAla 541
266 CTCCAATGCTGGCTAGAGCCACCGGACGAGCGCTAGTCCAATCTTG 217

```

```

541 hrProAlaLeuIlePheAlaAlaThrAlaAlaAlaThrProAlaPhe 557
    |||||
216 AGAACCGCTGCCAG.....GCCGAGCCGTACTGCATCGT 179
    |||||
558 AsnPheAlaAlaThrAlaAlaThrProAlaThrAlaAlaThrProAla 574
178 GAGATCCACCGGACGAGCGCGTAGCTCCAACTTGTGCTACCGCTGCC 130
    |||||
574 uIlePheAlaAlaThrAlaAlaThrAlaAlaThrProAlaThrAlaAla 591
129 .....AGGCGCAGCGCTTACTCCATGCTGCT..... 100
591 heAsnPheAlaThrAlaAlaThrAlaAlaThrProAlaThrAlaAlaPhe 607
99 .....AGAGCCACCGGACGACGCGCTACTGCATCTG 65
    |||||
608 AsnPheAlaAlaThrAlaAlaThrProAlaThrAlaAlaThrProAla 624
64 GAGAACCGCTGCCAGGCGCAGCGCTAGCT 34
    |||||
624 uMetPheAlaAlaThrAlaAlaThrProAla 634

```

seq\_name: sp\_bacteria:09SOT3

```

seq_documentation_block:
ID 09SOT3 PRELIMINARY: PRT: 565 AA.
AC 09SOT3:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE FLAGELLIN.
GN FLIC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U1A-44;
RA Ohnishi K., Ishioka K., Matsuda T., Harayama S.;
RT "Cloning of H antigen genes in E.coli serotypes and expression in
RT E.coli K-12."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB028480; BAA85089.1;
DR INTERPRO: IPR001029;
DR INTERPRO: IPR001492;
DR PFAM: PF00669; Flagellin_N; 1.
DR PFAM: PF00700; Flagellin_C; 1.
DR PRINTS: PR00207; FLAGELLIN.
SQ SEQUENCE 565 AA; 58528 MW; 68EFB39B6436342 CRC64;

```

```

alignment_scores:
  Quality: 170.00 Length: 543
  Ratio: 0.639 Gaps: 25
Percent Similarity: 48.987 Percent Identity: 22.284

```

alignment\_block:

US-09-306-689-12 x 09SOT3

Align seg 1/1 to: 09SOT3 from: 1 to: 565

```

469 CGCAATATATTGCAACAGCTCAACAGCTTGAAGCAAGATCAACCGC 518
    |||||
15 GlnAsnSnIleAsnLysAsnGlnSerAlaLeu.....SerSerSe 28
519 TATTGGCTTAACTGAGCGTGGCATTTGTATTCGCTCCACAAATTTATA 568
    |||||
28 rIleGluArgLeuSerGlyLeuArgIleAsnSerAlaLysAspSp. 44
569 AATTGTACAGAAAATTAAGCAGGCCAAGCATTA..... 603
    |||||

```

```

1204 ACTTTAGAGAGTTAATGCGGACGGTTTAAAAAATATAGCATATACGAGAC 1253
1205 ++++++::: ++++++:::
351 SerThraSnglyGlyGlyValLysPheLysAspThr...ValSerSerAs 366
1254 TAATTATTATAGCAGAAATATCAGCGGGAGACAGGAGACTATTGATGATCGG 1303
1255 ++++++::: ++++++:::
366 palLeuLeuAlaGlnValLysAlaAspSerThrAlaAsnValLysL 383
1304 TTACTGCATTAATACCGCATG.....GCCGATATGCTGGTGT 1344
1305 ++++++::: ++++++:::
383 LerThpheaSnaSnglyProLeuSerPheThrAlaSerPheGlnaSngly 399
1345 GTGTGTGCTGCTGCAGCCGATTAACATTTGAAAAGATTAAACATTAATCT 1394
1346 ++++++::: ++++++:::
400 ValSerGlySerAlaAlaSer.....AsnAl 408
1395 TGTCATCAGCAATAGCAAAAAAGAGAAAGTACCATTCAAACTGGTTC 1444
1396 ++++++::: ++++++:::
408 aLaIyrrLeaSpSerGlnGlyGluLeuThrThrThrGluSerTyr.... 423
1445 GAGAGCGCTGATTTTGTCTAAGACAGTGGCTTAATTTAAGCACTAAAGAT 1494
1446 ++++++::: ++++++:::
424 .....AsnThrSndYrSerValaSpLysAsp 432
1495 GAGAAATTCAGAGAAATCATCGTTCAAATGGCGAGGAGTACCTCCAA 1544
1496 ++++++::: ++++++:::
433 ThrGlyAlaValSerValThrGlyGlySerGly.....ThrIly 446
1545 GCAGCTGATGATCTTATTCGCAAA.....GGTAACGGCAAA 1582
1546 ++++++::: ++++++:::
446 sYzrAlaAlaAsnValGlyAlaGlnAlaIyrrValGlyAlaAspGlyL 463
1583 TTACCAAGATGACCTACAAAGTTTGATACATCATGATTCGCACAA 1632
1584 ++++++::: ++++++:::
463 eUrThrThraSnrThrSerThrGly.....Ser 472
1633 CATACCAAAATGTGACAAACAGCTTGATTAAGTATCATCTGATAG 1682
1634 ++++++::: ++++++:::
473 AlaThrIySaSpProLeuAsnAlaLeuAspGlnAlaAlaSerIleAs 489
1683 TGCAATTTACTCGTCTAATGATTCGAGAAATGATTAGTGGCTCCAACTT 1732
1684 ++++++::: ++++++:::
489 pLysPheArgSerSer.....LeuGlyAlaAlaGln 500
1733 CAATGTTGATCAAAAGTTTATCTCTCTT 1761
1734 ++++++::: ++++++:::
500 sNaIrgLeuAspSerAlaValIyrrAsnLeu 509

seq_name: sp_bacteria:Q9RY75

seq_documentation_block:
ID Q9RY75 PRELIMINARY: PRT: 1467 AA.
AC Q9RY75;
DT 01-MAY-2000 (TREMBLrel. 13, created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR0075.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
ON NCBI_TaxID=1299;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Elsen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffitt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;

```

RT "Genome Sequence of the Radioresistant Bacterium Deinococcus  
RT radiodurans R1.";  
RL Science 286:1571-1577(1999).  
DR EMBL: AE001870; AAF09665.1; -  
DR TIGR: DR0075; -  
DR INTERPRO: IPR00122; -  
SQ SEQUENCE 1467 AA; 148507 MW; 8964BA26A82693FF CRC64;

## alignment\_scores:

Quality: 169.00 Length: 838  
Ratio: 0.426 Gaps: 40  
Percent Similarity: 47.375 Percent Identity: 22.196

## alignment\_block:

us-09-306-689-12 x Q9RY75 ..

Align seg 1/1 to: Q9RY75 from: 1 to: 1467

```
43 CTGGCGCCCTGGCAGCGCTTCT...CAAGATTGAGCTACGGCTGCGTCC 89
||||| ||||| ||||| ||||| |||||
194 LeuArgAspGlyAlaThrAspLeuGlnGlnGlyThrValLysLeuArgAs 210
90 GGGTGGCTCT.....AGCCAGCATTTGGAGCT 115
||||| ||||| ||||| ||||| |||||
210 pGlyAlaSerAspLeuAlaAspGlyAlaAsnArgAlaAlaAspGlySerA 227
116 ACGGCTCGCCGCTGGCAGCTAGCCAGATTGGAGCTACGGCTGCGT 165
||||| ||||| ||||| ||||| |||||
227 rGlySerLeuAlaAspGlyAlaGlyThrLeuSerSerGlyValGlyArgLeu 243
166 CCGGCGATCTCAGCATTTGAGCTACGCGCTGCGCTGCGAGCGATT 215
||||| ||||| ||||| ||||| |||||
244 ThrAspGlyThrValLysLeuSerAspGlyLeuArgGlnLeuGluAlaAl 260
216 TCAAGATTGGAGCTACGCGCTGCGTCCG....GGTGGCTTACCGAC 259
||||| ||||| ||||| ||||| |||||
260 aaLProGlyGlnThrGlnLeuAlaProLeuArgSerGlyAlaGlyGln 276
260 ATTGAGACTACGCGCTGCGCTGCGTCCG..... 288
277 .....LeuArgThrGlySerGlnSerLeuAlaSerGlyLeu 288
289 .....GGTAGCCAAAGATTGAGCTACGCGCTGCGTCCG.. 321
||||| ||||| ||||| ||||| |||||
289 GlyGlnLeuAlaAspGlyThrAspGlnLeuAlaAlaGlyAlaLysAl 305
322 ....GGTGGATTCTAGCTTCCCAAAACTGGGCAAAAAAATTATCTCT 367
||||| ||||| ||||| ||||| |||||
305 aaSnGlyGlyAlaGlnGlnValAlaAlaGlyAlaSnLysGlnLeuAlaAlaG 322
368 ATTATCCCAAAATTACCAATATGATTAACAAGTAATGATTAGCT 417
||||| ||||| ||||| ||||| |||||
322 InLeuProGln.....LeuGlnSerGlyLeuThr 331
418 GATTATGCAAAAGCGCCGAGAGTTGGGATTGAGGTACAAAGAGA 467
||||| ||||| ||||| ||||| |||||
332 GlnLeuSerSerGlyAlaAspLysLeuAla..... 341
468 ACGCAATATATTGCAACAGCTCAACAGCTTGGCAGCATTAACCG 517
||||| ||||| ||||| ||||| |||||
342 .....AlaGlyAlaLysSerAlaAsnGlyGlyAlaGlnGln 354
518 CTATTGGC.....TTAAGTGGCT..... 537
||||| ||||| ||||| ||||| |||||
354 aAlaAlaGlyGlyAlaSnLysGlnLeuAlaGlnGlnLeuProLysLeuGlnGly 370
538 GGCATTGTGTATCCGCTCCAAATGATTAATGCTACAGAAAACATA 587
||||| ||||| ||||| ||||| |||||
371 GlyLeuThrGlnLeuSerGlnGlyAlaAspLysLeuAlaAlaGlyAlaLys 387
588 AGCAGGCCAAGCATTTAGTCTGCGGAAGCATTTGACAAAATGCAATA 637
||||| ||||| ||||| ||||| |||||
```

```
387 sSer.....AlaAsnGlyGlyAlaGlnGlnValAlaGlyLysAlaLysG 402
638 AAGCCAAATGTAATTATCTGCATTCATCTATTTAGCGCTGATATG 687
||||| ||||| ||||| ||||| |||||
402 InLeuAlaAspGlnLeuProGlnLeuGlnSerGlyLeuThrGlnLeuSer 418
688 GCTGAATGATTTAGATGAGCCTTACGAATATAACAGCAACAACATGC 737
||||| ||||| ||||| ||||| |||||
419 GlnGlyAlaAspLysLeuAlaAlaGlyAlaLysSerAlaAsnGlyAla 435
738 T.....CTTGCTAAAGCTGGCTGGAGCTACAAATTCATTA..... 774
||||| ||||| ||||| ||||| |||||
435 aGlnGlnValAlaAlaGlyAlaSnLysGlnLeuAlaAspGlnLeuProGln 452
775 .....ATTGAATAATATTGCTAAT 792
452 eugGlnSerGlyLeuGlnGlnLeuSerGlnGlyAlaAspArgLeuAlaSer 468
793 TCAGTAAACACCTTGACGAATTTGGTACGCAATTAAGT..... 831
||||| ||||| ||||| ||||| |||||
469 GlyAlaLysAspAlaAsnAlaGlyAlaGlnGlnValAlaAlaGlyAlaSn 485
832 CAATTGGTTCAAACTACAAATATCAAA...GGCTTAGGACTTTAG 877
||||| ||||| ||||| ||||| |||||
485 sGlnLeuAlaAlaGlnLeuProGlnLeuGlnSerGlyLeuGlnLeu 502
878 GAGCAAAACCTCAAAATATC...GGTGGACTGTGATTAAGCTGGCTGCT 924
||||| ||||| ||||| ||||| |||||
502 eArgGlyAlaAspArgLeuAlaSerGlyAlaLysGlyAlaAsnAlaGly 518
925 TTAGATGTTATCTCAGG..... 942
519 AlaGlnGlnValAlaAlaGlyAlaSnLysGlnLeuAlaThrGlnLeuProG 535
943 CTATTATCGGGCGCCAGCAGCTGCATCTGTACGATTAATATGCTT 991
||||| ||||| ||||| ||||| |||||
535 nLeuGlnSerGlyLeuGlnGlnLeuSerGlnGlyAlaAspArgLeuAla 552
992 CAACAGCTAAAAAAGTGGTGGCTGGGTTT..... 1020
||||| ||||| ||||| ||||| |||||
552 eArgGlyAlaLysAspAlaAsnAlaGlyAlaGlnGlnValAlaAlaGlyAla 568
1021 ...GAATTGGCAAAACCAAGTT.....GTGGTAATAT 1049
||||| ||||| ||||| ||||| |||||
569 LysGlnLeuAlaThrGlnLeuProGlnLeuGlnSerGlyLeuGlnLe 585
1050 TACCAAGCCGTTCTTCTTACATTAGCCCAACGCTTGCACAGCTT 1099
||||| ||||| ||||| ||||| |||||
585 uSerGlnGly.....AlaAspArgLeuAlaSerGlyA 596
1100 TATCTTCAACTGGGCTGCTGCTTTAATTGCTTACTGTTCTTCTT 1149
||||| ||||| ||||| ||||| |||||
596 IaAsnAspAlaAsnSerGlyAlaLysGlnLeuAla..... 607
1150 GCGATTAGCCCATTTAGCATTTGCCGT.....ATTGCCGATAAATT 1190
||||| ||||| ||||| ||||| |||||
608 .....AlaGlyAlaSnLysGlnLeuAlaAspGlnLe 617
1191 TATATGCAAAAAGT.....TTAGAGAGTTATGCCGACGCT 1228
||||| ||||| ||||| ||||| |||||
617 uProGlnLeuGlnSerGlyLeuGlnGlnLeuSerGlnGlyAlaAspArgL 634
1229 TTAATAAATAGCTATGAC..... 1248
||||| ||||| ||||| ||||| |||||
634 eAlaLysSerGlyAlaAsnAspAlaAsnAlaGlyAlaLysGlnLeuAlaAla 650
1249 GGAGATAATTTATTAGCA.....GAATATCAGCGCGGAGACAGG 1286
||||| ||||| ||||| ||||| |||||
651 GlyAsnArgGlnLeuAlaThrGlnLeuProGlnLeuGlnSerGlyLeuGln 667
1287 GACTATTGATCATCGGTTACTGCAATTAATACCGATGGCCGCTATTG 1336
||||| ||||| ||||| ||||| |||||
667 nGlnAlaSerAlaGlyAlaAspArgLeuAsnThrGlyAlaAlaGlnLeuAla 684
```

```

1337 CTGGTGTGTGCTGCTGCTGAGCCGATTACATTGTAAGTTAA 1386
||| |||
684 lathcylthrlleuInleuAsn.....AspAlaVallys 655
1387 CATTAATCTTGATCATCAGATACGAAAAAGAGAAAGTGCATTCAGAA 1436
||||| |||
696 AsnGlnAlaLeuLeuProGlyInleuLysGlnGlyValAlaLysValas 712
1437 CTGGTCCAGAGAGGCTGATTGCTTAAGACGTGCTTAATTATAAGCAA 1486
||| |||
712 n.....SerGlyAlaMetGluValHisSerGlyThrAspT 724
1487 CTAAAGATGAGAAATCGAAGAAATCGGTCGCAAAATGGCGCGGATC 1536
||| |||
724 hleuSerAsnAlaLeuProGlnLeuGlnAlaGlyAlaGlyAlaVal 740
1537 ACCTCA.....AAGCAAGTTGATGATCTTATCGCAAAAGGTAA 1574
||||| |||
741 ThrGlyAlaSerGlnLeuAlaGlnGlyAlaAspLysLeuAlaAlaGlyTh 757
1575 CGGCAAAATTTACCCAGATGAGCTATCAAAAGTTGTGATACATGAAT 1624
||||| |||
757 rAlaGlnLeu..... 760
1625 TGCTCAACATAGCAAAATGCAACAGAGCTAGATAAGTTA..... 1668
||||| |||
761 .....SerGlnGlyAlaThrThrLeuSerSerLysLysGlu 773
1669 .....ATCTCATCTGTAGTGCATTTACCTCGCTCATGA 1703
||||| |||
774 AlaGlnGlnGlySerThrAlaAlaValGlnGlyValGlnGlnLeuSerG 790
1704 TTCGAGAAATGATTTATGTTGCTCAACTTCATGTTGATCAA..... 1746
||||| |||
790 uGlyAlaThrArgLeuAlaAlaGlyThrAlaGlnLeuSerGlnGlyAla 807
1747 ..AGTTATCTCT..... 1758
||||| |||
807 alThrLeuSerLysLysGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 823
1759 CTTCATTTGCTAGGAGATCTGAGATGAGAGCTGAGGCTGCGCC..... 1804
||||| |||
824 ValGlnGlyAlaGlnGlnLeuGlnGlyLysLysLysLysLysLysLys 840
1805 .....CTGCGAGCGGCTTCAAG 1822
||||| |||
840 lYThrAlaGlnLeuSerGlnGlyAlaAsnThrLeuSerSerLysLys 856
1823 .....ATTGAGCTAGCGGCTGCGGCTGCGCTCTGACCGACAT 1863
||||| |||
857 GluAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 872
1864 TGGAGCTAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 1891
||||| |||
872 uGlyGlnGlyAlaAspLysLeuAlaAlaGlyThrAlaGlnLeuSerGln 889
1892 .....GTAGCCCAAGATTGAGAGCTACG 1912
||||| |||
889 lYAlaThrThrLeuSerSerLysLysGlnAlaGlnGlnGlnGlnGln 905
1913 GGCTGCTGCGGCTGATCTGACGATTCAGCATTCAGGCTGCGGCTGCG 1962
||||| |||
906 AlaAlaValaGlnGlyAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 922
1963 A.....GGGTTCTCAAGATTGAGCTAGCGGCTGCGGCTGCGGCTG 2003
||||| |||
922 AlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 938
2004 CTCTAGCCAGCATTTGA.....GCTAGCGGCTGCGGCTGCGGCTGAGCC 2050
||||| |||
938 yslLeuysGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 954

```

```

2051 AAGATTGAGACT 2062
||||| |||
955 LysLeuAlaAla 958

```

```

seq_name: sp_bacteria:Q9PC04

```

```

seq_documentation_block:

```

```

ID Q9PC04 PRELIMINARY; PRT: 1190 AA.

```

```

AC Q9PC04:

```

```

DT 01-OCT-2000 (TRENBLREL. 15, Created)

```

```

DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)

```

```

DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)

```

```

DE SURFACE PROTEIN.

```

```

GN xfp1981.

```

```

OS Xylella fastidiosa.

```

```

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

```

```

OC Xylella.

```

```

OX NCBI_TaxID=2371;

```

```

RP [1]

```

```

RC SEQUENCE FROM N.A.

```

```

RX STRAIN=9A5C; PubMed=10910347;

```

```

RX MEDLINE=20365717;

```

```

RA Simpson A.J.G., Rehnach F.C., Arruda P., Abreu F.A., Acencio M.,

```

```

RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

```

```

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

```

```

RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,

```

```

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

```

```

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfi H.,

```

```

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

```

```

RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,

```

```

RA Garnier L., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

```

```

RA Ho P.L., Hohetsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

```

```

RA Krieger J.E., Kurame E.E., Laigret F., Lambais M.R., Leite L.C.C.,

```

```

RA Machado E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

```

```

RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Marino C.L.,

```

```

RA Marques C.F.M., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,

```

```

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

```

```

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Neto L.E.S.,

```

```

RA Nhani A.J.R., Nobrega F.G., Nunes L.R., Oliveira M.A.,

```

```

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

```

```

RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,

```

```

RA Queiglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

```

```

RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

```

```

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,

```

```

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

```

```

RA de Souza A.P., Terenzi M.F., Trufi D., Tsai S.M., Tsuchino M.H.,

```

```

RA Zago M.A., Zatz M., Meidams J., Setubal J.C.;

```

```

RT "rpe genome sequence of the plant pathogen Xylella fastidiosa.";

```

```

RL Nature 406:151-157(2000).

```

```

DR EMBL; AE004017; AAF84783.1;

```

```

SQ SEQUENCE 1190 AA; 11846 MW; 756741B0C8D787CC CRC64;

```

```


```

```

alignment_scores:

```

```

Quality: 168.50 Length: 847

```

```

Ratio: 0.428 Gaps: 40

```

```

Percent Similarity: 46.517 Percent Identity: 21.842

```

```

alignment_block:

```

```

US-09-306-689-12 x Q9PC04 ..

```

```

Align seg 1/1 to: Q9PC04 from: 1 to: 1190

```

```

40 GGCTGCTGCGGCTGAGCGGTTCTCAAGATTGAGCTAGCGGCTGCGTCC 89

```

```

155 GlyLeuValAspLysSerGlyAla.....ThrTyrSerMetArgLys 168

```

```

90 GGGT.....GGCTTAGCCAGCATTTGAGCTAGCGGCTGCGGCC 121

```

```

168 uGlySerIleAlaThrMetAsnGlySerAlaGlnGlnGlnGlnGlnGln 185

```

```

122 TGGCGCCCTGCGGCTGAGCCAA.....GATTGAGCATGAGCGGCTG 162

```

[illegible]

456	laValasnhgclglnleuGlnAlaValasrApThrlAlaserLyscl	472
925	TTAGATGTTATC.....TCAGGCGATTATATGGGGCC	956
473	TrpAsnleuAlaserGlyThrAsnSerSerAsnValAlaProGlyAl	489
957	AACAGCTGCACCTTGACTGCAGATMAAAATGCTTCACACCTATAAAAG	1006
489	asrValAspLeuLysAsnThrlAspGlyAsnleuLeuThrLysAlar	506
1007	TGGGTGGGGT.....TTGAAATGGCAACCAATT.....	1038
506	IeLylLLeAsnAspAlThrlrPheAsnleuAlaThrlAlaLeuGlnAlaasp	522
1039	.....GTTGGTAATATATACCAAGCCGTTCTCTTACATTTAGC	1079
523	serLeuThrlThrglyAsnThrlaMetThrlrAspGlyAlThrlValgl	539
1080	CCACGCTGTTCAGCAGGTTTATCTTCACATGCGGCTGTGGCTGTTAA	1129
539	YserAsnValThrlLeuGly.....SerThrlcylLeuValIleThrlAspG	554
1130	TTGCTCTACTGTTTCTCTTGGCATTC.....CCA	1161
554	lyProserValThrlserSerGlyIleSerAlaGlyAsnGlnLysIleThr	570
1162	TTAGCATTTCCGCGATTGGCCGAT.....AAATTTAATCATGC	1199
571	AsnValAlaAlaGlyThrlrAlaAspThrlrAspAlaValAsnPheSerGlnLe	587
1200	AAAAATTAGAGAGTTATGCCGACCGCTTMAAAATATAGCATATGACG	1249
587	uGlnAlaValSerSerThrlAlaSerLys.....G	597
1250	GAGATATTTATTATAGCAGATATACAGCGGGCA.....	1281
597	lyTrpAsnleuLeuAlaSerGlyAlaAsnSerSerAsnValAlaProGly	613
1282	.....ACAGGAC	1289
614	GlySerValAspLeuLysAsnSerAspGlyAsnleuLeuIleThrlLysTh	630
1290	TATGATGATGGGTTACGCAATTAATACCCGATTTGGCCGCTATTGCTG	1339
630	ThrlrAspSerAsnAspValThrlrPheAsnleuAlaThrlAlaLeuLysVala	647
1340	GNGGTGTGCTGCTGCTGCAGCCGATTAACATTGAAAAAGTT.....	1383
647	spSerLeuThrlThrlGlyAsnThrlaMetThrlThrlAspGlyAlThrlVal	663
1384	.....AACATAATCTTGTCATCAG.....	1404
664	GlySerAsnValThrlLeuGlySerThrlGlyLeuValIleThrlAspGlyPr	680
1405	.....AATAGCAAAAAAGAGAAAGTGCACATTC	1432
680	oSerValThrlrSerSerGlyIleSerAlaGlyAsnGlnLysIleThr...	695
1433	AAACATGTTCCGAGAGCTGATTTGTCMAAAGAGTGCCTAATTATATAA	1482
696	.....AsnValAlaAlaGlyThrlAlaAspThrlAsp	705
1483	GCAACTMAAGATGAGAAATTCAGAGAAATC.....	1512
706	AlaValAsnPheSerGlnleuGlnAlaValSerSerThrlAlaSerLysGln	722
1513	.....ATCGTCAAAAT.....GGCG	1528
722	ylTrpAsnleuLeuAlaSerGlyAlaAsnSerSerAsnValAlaProGlyG	739
1529	AGCGATTCACCTCAAAAGCAAGTTGATGAT.....CTTATGCGCAAAAGCT	1572
739	IleValAspLeuLysAsnThrlrAspGlyAsnIleValIleSerLysGln	755

```

1573 AACGGCAAAATTCACCAAGATGAGCTATCAACAAATGTTGGTAACTCTAGA 1622
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
756 SerGly.....SerAspValLeuPheAsnLeuSerSerLeuLy      770
    :|||
1623 ATTG..... 1626
770 sLeuAspPysLeuThrValGlyAspThrValMetThrThrAsnGlyValT 787
    .....CTCAAAcAT 1655
1627 .....
787 hrValGlySerGlyValThrLeuGlySerMetGlyLeuValIleThrAsp 803
    1636 AGCAAAATGTCAC.....AACGCTTAGATAGTAATGATCTC 1673
    :|||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:
804 GlyProSerValThrSerSerGlyIleAsnMetGlySerGlnIleIleH 820
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1674 ATCTGTAAAGTCATTTACCTCGCTCAATGATTCAGAAATGATTTAACTGG 1723
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
820 rAsnValAlaIleGlyThrAlaAspThrAlaAspThrAlaValAlaLeuSerGlnL 837
1724 CTCACACTTCATGTGTGATCAAGATTATCTCTCTTCATTTGGCTAGG 1773
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
837 euAsnThrAlaMetAlaGlySerGlyAlaIalysSerValHistyr..... 851
1774 GGATGTCAGCATTTGGAGCTACGGCTCGGCCCTGACAGCGGTCTCAGA 1823
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
852 ..TyrSerThrThrAspGlyGlyThrGlnGlnGlyAlaAsnThrAsnGlyAs 867
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1824 TTGAGAGCTACGCGCTCGCTCGGGGTGCTCAGACAGCATTTGAGACTAGC 1873
    P:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
867 P:GlyAlaThrGlyThrArgSerIleAlaVal...GlyValGlyThrLeu 882
1874 GCCTCGCGCCCTGGCAGCGGTAGCCAAAGATTGAGACTACGGCTCGCTCCG 1923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
883 AlaSerAla.....GlnGlyValaThrAla..... 890
1924 GGTGGATTCACATTTGAGAGCTACGGCTCGGCCCTGACAGCGGTCTCTCA 1973
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
891 .....ValGlySerGlyAlaAlaAlaSerGlyLys..... 900
1974 AGATTGAGACTACGGCTCGCTCGGGGTGCTCAGACAGCATTTGAGACT 2023
    |||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:
901 .....GlySerThrAlaIleGlyArgAsnAlaValAlaSerAlaAspGly 915
2024 ACGCGCTCGCCCTGGCAGCGGTAGCCAAAGATTGAGACT 2062
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
916 SerVal...AlaLeuGlyAspGlyAlaLysAspGlyAla 927

seq_name: sp_bacteria:068729

seq_documentation_block:
ID 068729 PRELIMINARY: PRT: 1525 AA.
AC 068729;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DE 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
PT PROTEIN H11514 HOMOLOG.
GN Y1044 OR YPWT1.11C.
OS Yersinia pestis.
OG Plasmid pMT1, and Plasmid pMT-1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
ON NCBL_TaxID=632;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K1M;
RA Hu P., Elliott J., McCready P., Skowronski E., Ganes J.,
RA Kobayashi A., Carrano A.V., Brubaker R., Garcia E.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K1M10+;

```

RA MEDLINE=99043898; PubMed=9826348;  
RX Lindler L.E., Plano G.V., Burland V., Mayhew G.F., Blatner F.R.;  
RT "Complete DNA sequence and detailed analysis of the Yersinia pestis  
RT KIM5 plasmid encoding murine toxin and capsular antigen.";  
RL Infect. Immun. 66:5731-5742(1998).  
RN 131  
RP SEQUENCE FROM N.A.  
RC STRAIN-CO-92 BIOVAR ORIENTALIS;  
RC Baker S.G., Mungall K.;  
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ  
RN 141  
RP SEQUENCE FROM N.A.  
RC STRAIN-CO-92 BIOVAR ORIENTALIS;  
RA James K.O., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN 151  
RP SEQUENCE FROM N.A.  
RC STRAIN-CO-92 BIOVAR ORIENTALIS;  
RA Karlyshev A.V., Wren B.W.;;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF053947; AAC13170.1; -  
DR EMBL: AF074611; AAC82704.1; -  
DR EMBL: AL117211; CAB55193.1; -  
KW Plasmid, Hypothetical protein.  
SQ SEQUENCE 1525 AA; 162113 MW; F14C1CB0B1434027 CRC64;

```
alignment_scores:
  Quality: 166.50      Length: 8833
  Ratio: 0.386         Gaps: 44
Percent Similarity: 48.811  Percent Identity: 21.0655
```

alignment\_block:  
US-09-306-689-12 x 068729

Align seg 1/1 to: 068729 from: 1 to: 1525

```

23 CTCACAGATTGAGCTACAGCGCTCGCCCTCGACACCGATTCTCAAGATTGG 72
   :::::::::::::: :::::::::::::: ::::::::::::::
513 ILeSerLeuGlyAlaValValAlaAlaAlaMetValAlaPhe..... 525
73 AGCTACGGCCTCGCCGCGGTGGCTCTTACGACAGATTGAGACTACGCGC. 121
   :::::::::::::: :::::::::::::: ::::::::::::::
526 .....LysLeuMetArgAlaGlyIleAlaGlyValIleGlyThrIaG 540
122 .....TCGCGCCTCGACCGCTA 139
   :::::::::: ::
540 LylGlnMetValAsnThrPheThrIlyrIysMetSerThrValLeuGlnAlaPro 556
140 GCCAAGATTGAGCTACGCGCCTCGCTCCGCGGTGATCTCAGACTT.... 184
   :::::::::::::: :::::: ::::
557 PheAsnLeuGlyAlaThrAlaValThrArgPheAsnArgAlaAlaIaArgMe 573
164 ..... 184
573 LeIlyLeuAlaProIleProSerLeuIlePheAlaIleArgGlyAlaIleT 590
185 .....GGAGCTACGCGCC.....TCGCGCCTCGACGCGCTCTC 217
   :::::::::: :::::: ::::::
590 hrGlyLeuGlnGlyAlaPheAlaGlyLeuThrAlaPheIleAlaIaAsn 606
218 AAGATTGAGCTACGCGCCTCGCTC...CGGCTGGCTCTAGCCAGACTT.. 262
   :::::::::::::: :::: ::::::::::::::
607 ProIleGlyAlaAlaPheThrValAlaAlaThrValAlaValAlaGlyLeuI 622
263 .....GGAGCTACGCGCCTCGCGCCTGCGACCG 290
   :::::::::: :::::: ::::::
623 eThrTyMetThrMetLeuArgSerGluThrSerIysValValAspIuI 640
291 TAGCCAAGATT.....GGAGCTACGCGCCTCGCGGTGAGACTAGCT 334
   :::::::::: :::: ::::::::::::::
640 LeArgLysIleProGluIaLeuThrThrAlaValAspValGlnAlaGlnMetAla 656

```

335 TCCCA..... 339  
657 AAlaAgAlaAlaGlueGlulysGlnllectlnArGaspGlnAlaLe 673  
340 .AAACGGGGCAAAAAAATATTCCTATATTCGCCAAATACCAAT 388  
673 ulysThrGlyLuser.....Valnsnry...T 682  
389 ATGATCTGCAACAAGCTAATGCTTACAGATTTAGTCAAAAGCCGAA 438  
682 ySerThrAlaAlaGly.....ProValAlaValLysGlnLuserLys 695  
439 GAGTGGGAGATTGAGTGCACAAAGACACAGCAATATTTATTCACAGC 488  
696 GluVal.....ValGlulAlaArgLeuLysLysAsnGlnGluLty 709  
489 TCAAACAGATTAGGCACGATTCACAAACCGCTATAGTGCTTAACAGCTG 538  
709 rGluArGThrThrGlyThrMetAlaLeuGlnLyspGlyAlaValAlaLysA 726  
539 GCATGTGTTATCGCTCCACAAATGTGATAATTCCTACAGAAACTAA 588  
726 rGluValAlaLysGlnAla.....AlaGlnSerGlnlLeGluLyslLeArg 740  
589 GCAGCGCAAGCTTAGTTCCTGCCCAAGCACTGACACAAATGCAAATA 638  
741 AlaAspAsnGlnValPheSerAlaThrPheVal.....Lys 752  
639 ACCCAAA...ACTGATATATCGCATTCACATCTATTTAGGCTCAGAT 685  
752 sAlaArGlnGlnAlaAlaLeuAspLyslLeGlnLysLe..... 764  
686 TGGCTGGAATGATTTAGTAGAGCCCTTACAGAAATACAGCAACCAACAT 735  
765 .....AsnAspAspLysSerLeuSerAspAspGlnLysAsnLys 777  
736 GCTTTGCT.....AAAGCTGGCTTGGAGCTAC 764  
778 LeuLeuAlaProLeuArGlnLulysValAsnLysSerLtyrLeuGlnProAl 794  
765 AAATTCATTAATGTAATAATTTGTAATTCAGTAATAAACACTTGCGAAT 814  
794 acGlnLysLeuValLtyrAspLeu...SerSerArgLysAsnAlaThrGlu. 809  
815 TTGGTGACCAATTAAGTCATTTGGTGCACAACTGACAAATATACAAAGC 864  
810 .....LysGlnlLeuAlaHisPheSerAspMetLeuGlnLysAlaLysLys 824  
865 TTAGGAGCTTTAGGAGACAACACTCAAAAAATACGCTGACCTGTATAAAGC 914  
825 GlnGlyAsnThr...GlnGlnValAlaLnsLysLeuGlnGly..... 836  
915 TGGCCTTGTTAGATGATATGTCAGGGCTA.....TTATCGGCGC 955  
837 .....SerlLeArgGlyLtyrGlnGlnHisLeuGlnAlaVal 848  
956 CAACACCTGCACCTGTACTTGCAGATTAATAATGCTTCAACAGATAAAA 1005  
848 AlAlaGlnGlnLeuThrGlnAlaGlnPheGlnLysLysAspSerAlaLys 864  
1006 GTGGGTGGGGGTTTGAATGGCAACCAA.....GTTGTGGTATAT 1049  
865 ThrGlnLysGlyVal...MetSerAsnGlnGlyThrValLeuGlnLysGln 880  
1050 TACC.....AAAGCGGTTCTCTTACACNTTTAGGCC 1081  
880 yThrThrAspLysAlaAlaGlnLysAlaLeuAlaGlnlntyrMetArGAsnG 897  
1082 AACGTTCTCACACAGTTATCTTCACT..... 1110  
897 lMetAspSerAlaThrLtyrGlnAlaGlnlntyrLeuProAspGlyThrAlaMet 913  
1111 .....GGGCTT..... 1116

914 LeuAspPheGluGlyLysAspProIleIleGlyProLysGlnLeuLysTrpGln 930  
1116 ..... 1116  
930 nLeuAsnLeuGlnLysAlaSerSerAlaSerSerLeuGlnLysMetSerA 947  
1117 ..... GTGGCGCTGTTAATTGGCTTCACGCTTC 1146  
947 spGluGlnArgAlaIleAlaIleAlaIleAlaLeuThrLysAlaArgGluGln 963  
1147 ...CTTGCATTAGCCCATTTAGCATTTCCGCTATTCGCGAATTAATTAA 1193  
964 AspAlaIleAlaIleAlaGlnLysAlaIleGlyGlnArgTrpThrAlaAsnAlaSerGln 980  
1194 TCATGCAAAAAGTTTAGACAGAGTTATCCGAACGCTTTAAAAATTAGGCT 1243  
980 nArgAlaIleAlaArgLysGlnGlnLysAlaAlaGlnArgLysLeuAlaIleGlyT 997  
1244 ATGAC.....GGAGTAAATTTA.....TTAGCAGAA 1269  
997 yGlnLysAlaLeuAspLysAlaAspGlnLeuMetGlyGlnMetGlyGln 1013  
1270 TATCACGGGGAGAACAGGCGATTATGATCATTACGCTGCTACTGCAATTAATAC 1319  
1014 SerSerLysAlaThrValSerPheAspGlnSerLeuAlaGAspThrThrLys 1030  
1320 CGCATTTGGCCGCTATTTCGTGTGTGTG...TCTGCTGCTGCACCGCAT 1366  
1030 sSerLeuThrGlnLeuAlaAsnAlaValProAsnGlnPheIleThrGln 1047  
1367 TAACATTTAAAAAGTTAACTAATCTTGTTC..... 1398  
1047 LuMetLleAspLysAlaLysSerArgLeuAlaAspLeuAlaAsnAlaSer 1063  
1399 .....ATCCGAATATGCAAAAAGCAAGACCATTCGA 1433  
1064 AspAspTrpArgGlnMetPheAsnArgArgAsnValGlnGlnMetIleSe 1080  
1434 AAACGTGTTCCAGAGGCGTATTTTGCTAAAGACAGTCGCTAATTATAAG 1483  
1080 rTrpTrpAlaProGlnSerAspSerIleIleSerIle...GlyTrpLysP 1096  
1484 CAACATAAGATGAGAAATCGAAGAATATCCTGCTCAAAATGCGCAGCG 1533  
1096 rSerArgGlnGlnLysValAlaAspPheAsnAspThrTyrFAsnArgAsn 1112  
1534 ATCACC..... 1539  
1113 LeuLysAlaLeuMetAspLeuArgAspGlnAlaSerAspProLysIleVal 1129  
1540 .....TCAAGCAAGTTGATGATCTTATCCGAAGAAGTAAACGC 1578  
1129 AlaLeuTrpThrLysGlnIleAsnGlnLeuValAlaIleGlyAsnThrA 1146  
1579 .....AAAATTCCCAAGATGAGCTATCAAAAGTTGTTGATACAT 1620  
1146 AlaLeuLleLysGlnTrpArgTrpAlaArgGlnLysLeuAlaLeuGlnLysTr 1162  
1621 GAATTCGCTCAACATATGCAAAAATGTGACAAAACCTTAAGTAAGTTAT 1670  
1163 GlnAsnLeu....AlaGlnIleLeuGlnLysSerTrpSerAsnLeuPhe 1177  
1671 CTCATCTGTAAAG.....GCATTTA 1690  
1177 eSerAsnMetThrAspThrLeuThrAspPheValMetLysGlyLysLeuA 1194  
1691 CCTGCTCTAATGATTCGAGAAATGATTAGTGGCTCAACACTTCATGTTG 1740  
1194 sPheSerSerLeuAlaGlnSerIleLeuArgAspIleThrAsnMetVal 1210  
1741 GATCAAGAATTATCTTCTCTCAATTTGCTAGGCGATCTCAGCATTTGAG 1790



1211 .ValylsThrInleThrLeuProLeuMetLeuGlyMetGlyT 1227  
1791 CTAGGGCTGGGCC.....CTGGACGGGTCTTC 1819  
1227 hTThrAlaIaGlySerSerGlnSerGlyAsnLeuSerGlyValAla 1243  
1820 AAGATTGAGCTACGGCCTGCTCCGGGTGGCTTAGCC..... 1858  
1244 SerAlaValAlaAsnGlnGlyValArgMetAsnAlaValAsnGlyAspLy 1260  
1859 .AGCATTTGA...GCTACGGCCTGGCCCTGGACGGGTAGCCAGATTG 1904  
1260 sSerValGlyGlnAlaThrLyGlnThrSerSerSerValSerGlyLeuG 1277  
1905 GAGCTACGGCCTGCTCCGGGTGATCTCAGCATTTGAGCTACGGCCTGC 1954  
1277 LyGlnThrThr...GlnGlnThrThrSerAlaIleGlySerAlaThrAsn 1292  
1955 GCCCTG.....GCAGCGTCTCAAGAT 1977  
1293 AlaIleGlyAsnTrpValAsnGlyLeuPheThrSerThrGlnAlaLysAs 1309  
1978 TGGAGCTACGGCCTGCTCCGGGTGCTTAGCCAGCAT..... 2017  
1309 PALaGlnThrLysAlaValLysThrSerIlePheSerMetGlnAsnLeuS 1326  
2018 ..GGAGCTACGGCCTGCGCCCTGGCAGCGGTAGCCAGATTGAGCT 2062  
1326 eTSerValThrGlyAlaLeuSerAlaAlaPheAlaMetLeuGlyAla 1341

**THIS PAGE BLANK (USPTO)**

OM of: US-09-306-689-12 to: SwissProt\_39.\* out\_format : pfs  
 Date: Mar 2, 2001 10:51 AM

About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frame+np2.model -DEV=rlp  
 -Q/cgcn2.1/usprot.spool/us09306689/runat.02032001\_102830\_9659/app\_query.fasta.1.2389  
 -DB=SwissProt\_39 -QPM=fastan -SUFFIX=isp -GAPOP=12.000  
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000  
 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
 -FCAPOP=6.000 -FCAPEXT=7.000 -YCAPOP=10.000 -YCAPEXT=0.500  
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsun62  
 -TRANS=human0.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct  
 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=pfs  
 -NORM-ext -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09306689.@CGN1\_1.176 -NCRU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
 -MAIL -THREADS=1

## Search information block:

Query: US-09-306-689-12  
 Query length: 2088  
 Database: SwissProt\_39.\*  
 Database sequences: 88757  
 Search time (sec): 97.840000

## score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
SwissProt_39:CYAA_BORPE	212.00	136.29	0.0002	1706	P15318 Bordetella pertussis
SwissProt_39:CYAA_BORBR	206.50	131.11	0.0003	1705	Q57506 Bordetella bronchise
SwissProt_39:TALA_DICDI	166.50	166.71	0.0052	2491	P54633 dictyostella bronchise
SwissProt_39:K2C1_HUMAN	166.50	160.50	0.0394	643	P04264 homo sapiens (human)
SwissProt_39:K220_HUMAN	166.00	160.01	0.0419	638	P01546 homo sapiens (human)
SwissProt_39:K22E_HUMAN	166.00	160.01	0.0419	645	P35508 homo sapiens (human)
SwissProt_39:K1C1_HUMAN	163.00	157.44	0.0603	622	P35527 homo sapiens (human)
SwissProt_39:GUNE_CELFI	161.50	152.25	0.0699	1045	P26225 cellucomonas finl.
SwissProt_39:PALI_MOUSE	159.00	143.41	0.0893	2541	P26039 mus musculus (mouse)
SwissProt_39:FLIC_SALON	157.50	153.75	0.1188	507	Q06974 salmonella oranienber
SwissProt_39:FLIO_YEAST	155.50	146.95	0.1449	995	P40642 saccharomyces cerevis
SwissProt_39:PALI_HUMAN	155.50	140.11	0.1363	2541	Q09490 homo sapiens (human)
SwissProt_39:SM41_HEMPU	154.00	152.05	0.1841	407	Q26264 hemocytotus pulcher
SwissProt_39:CPN1_DROME	153.00	145.61	0.1978	865	Q02910 drosophila melanogaste
SwissProt_39:FLA3_CAMEL	152.50	148.13	0.2158	574	Q46113 campylobacter jejuni.
SwissProt_39:SLAP_CAMEL	150.50	142.70	0.2663	933	P35827 campylobacter fetus.
SwissProt_39:ANP_NOTCO	145.50	139.20	0.4928	790	P24856 notiohenia coriiceps
SwissProt_39:APMD_PIG	145.00	135.99	0.5108	1150	P12021 sus scrofa (pig)
SwissProt_39:NL6_YEAST	144.50	135.76	0.5438	1113	Q002630 saccharomyces cerevi
SwissProt_39:K1C1_HUMAN	144.00	139.88	0.6019	593	P13645 homo sapiens (human)
SwissProt_39:SLAP_CAVCR	144.00	135.88	0.5808	1025	P35828 caluolacter crescent
SwissProt_39:MYH_YEAST	143.50	133.32	0.6056	1367	P08640 saccharomyces cerevi
SwissProt_39:RBPI_PLYAV	143.50	137.91	0.5771	2869	Q00798 plasmodium vivax (st
SwissProt_39:FLJB_SALAE	142.00	139.23	0.7752	500	P52615 salmonella abortus-ed
SwissProt_39:SVS2_RAT	141.00	139.67	0.8856	414	P22006 rattus norvegicus (rat)
SwissProt_39:190K_RICRI	140.50	136.86	0.8427	2249	P15921 rickettsia rickettsi
SwissProt_39:MCPA_CAVCR	140.00	135.36	0.9639	657	Q00986 caluolacter crescent
SwissProt_39:PI1J_PSEAE	140.00	135.09	0.9675	682	P42257 pseudomonas aeruginos
SwissProt_39:FLIC_SALAE	139.00	136.78	1.05	507	Q06968 salmonella berta. fld
SwissProt_39:MYSN_AACCA	139.00	128.35	1.04	1509	P05659 acanthamoeba castelli
SwissProt_39:G33_BP14	138.50	129.03	1.11	1289	P18771 bacteriophage t4. lat

seq\_name: SwissProt\_39:CYAA\_BORPE

## seq\_documentation\_block:

ID CYAA\_BORPE STANDARD; PRT; 1706 AA.  
 AC P15318;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE BIFUNCTIONAL HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (CYCLOLYSIN) (ACT)  
 DE (AC-HLY) [CONTAINS: CALMODULIN-SENSITIVE ADENYLATE CYCLASE  
 DE (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (ADENYLATE CYCLASE); HEMOLYSIN].  
 GN CTA OR CTA.  
 OS Bordetella pertussis.  
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
 CC Bordetella.  
 CC [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-18323.  
 RX MEDLINE=88216178; PubMed=2897067;  
 RA Glaser P., Ladant D., Sezer O., Pichot F., Ullmann A., Danchin A.;  
 RT "The calmodulin-sensitive adenylyate cyclase of Bordetella pertussis:  
 RT cloning and expression in Escherichia coli.";  
 RL Mol. Microbiol. 2:19-30(1988).  
 RN [2]  
 RP SEQUENCE OF 1489-1706 FROM N.A., AND BIFUNCTIONAL PROTEIN DESCRIPTION.  
 RC STRAIN-18323;  
 RX MEDLINE=89091151; PubMed=2905265;  
 RA Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;  
 RT "Secretion of cyclolysin, the calmodulin-sensitive adenylyate cyclase-  
 RT haemolysin bifunctional protein of Bordetella pertussis.";  
 RL EMBO J. 7:3997-4004(1988).  
 RN [3]  
 RP DOMAINS.  
 RX MEDLINE=91177021; PubMed=2007407;  
 RA Munier H., Gilles A.-M., Glaser P., Danchin A., Sarfati R., Barzu O.;  
 RT "Isolation and characterization of catalytic and calmodulin-binding  
 RT domains of Bordetella pertussis adenylyate cyclase.";  
 RL Eur. J. Biochem. 196:469-474(1991).  
 RN [4]  
 RP MUTAGENESIS.  
 RX MEDLINE=89251630; PubMed=2542030;  
 RA Glaser P., Elmogolou-Lazaridou A., Krin E., Ladant D., Barzu O.,  
 RA Danchin A.;  
 RT "Identification of residues essential for catalysis and binding of  
 RT calmodulin in Bordetella pertussis adenylyate cyclase by site-directed  
 RT mutagenesis.";  
 RL EMBO J. 8:967-972(1989).  
 RN [5]  
 RP MUTAGENESIS.  
 RX MEDLINE=91266896; PubMed=2050107;  
 RA Glaser P., Munier H., Gilles A.-M., Krin E., Porumb T., Barzu O.,  
 RA Sarfati R., Pellecier C., Danchin A.;  
 RT "Functional consequences of single amino acid substitutions in  
 RT calmodulin-activated adenylyate cyclase of Bordetella pertussis.";  
 RL EMBO J. 10:1683-1688(1991).  
 RN [6]  
 RP REVIEW.  
 RX MEDLINE=93119764; PubMed=8418825;  
 RA Danchin A.;  
 RT "Phylogeny of adenylyl cyclases.";  
 RT Adv. Second Messenger Phosphoprotein Res. 27:109-162(1993).  
 RN [7]  
 RP PALMITOYLATION AT LYS-983.  
 RX MEDLINE=95025937; PubMed=7939682;  
 RA Hackett M., Guo L., Shabanowitz J., Hunt D.F., Hewlett E.L.;  
 RT "Internal lysine palmitoylation in adenylyate cyclase toxin from  
 RT Bordetella pertussis.";  
 RL Science 266:433-435(1994).  
 RN [8]  
 RP PALMITOYLATION AT LYS-860.  
 RX MEDLINE=99214144; PubMed=10196151;  
 RA Basar T., Havlicek V., Bezouskova S., Halada P., Hackett M., Sebo P.;  
 RT "The conserved lysine 860 in the additional fatty-acylation site of  
 RT Bordetella pertussis adenylyate cyclase is crucial for toxin function  
 RT independently of its acylation status.";  
 RL J. Biol. Chem. 274:10777-10783(1999).  
 CC -1- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF  
 CC BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN  
 CC CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL  
 CC CELL FUNCTION.  
 CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.



```

926 TAGATGTATCTCAGGGCTATATCGGGCAACAGCTGACTTACTT 975
    : : : : : : : : : : : : : : : : : : : : : :
539 eAlaLeuGlyGlyGlyLeuAlaAlaAlaValGlyValMetSerLeu 555
    : : : : : : : : : : : : : : : : : : : : : :
976 GCAGATTAATAATGCTTCAACAGCTTAATAAGTGGGCGGTTTGAAAT 1025
    : : : : : : : : : : : : : : : : : : : : : :
556 ThrAsp...AspAlaProAlaGlyGlnLysAlaAlaAlaGlyValGlu 571
    : : : : : : : : : : : : : : : : : : : : : :
1026 GGCACCAACAGTTGTTGTAATATTAACAACCGCTTCTCTTACATTT 1075
    : : : : : : : : : : : : : : : : : : : : : :
571 eAlaLeuGlnLeuThrGlyGlyThrValGlnLeuAlaSerSerLeuAla 588
    : : : : : : : : : : : : : : : : : : : : : :
1076 TA.....GCCAAGCTGTGGCAGCGAGTTATCTTCACTGGG 1113
    : : : : : : : : : : : : : : : : : : : : : :
588 eAlaLeuAlaAlaAlaArgGlyValThrSerGlyLeuGlnValAlaGly 604
    : : : : : : : : : : : : : : : : : : : : : :
1114 CCGTGGCGCTTAAATGCTTACTGTTTCTGCGATTAAGCCCATTT 1163
    : : : : : : : : : : : : : : : : : : : : : :
605 AlaSerAlaGlyAlaAlaAlaGlyAlaLeuAlaAlaLeuSerProMet 621
    : : : : : : : : : : : : : : : : : : : : : :
1164 ACCATTTCGGCGTATTCGGATTAATTAATCATGCAAAAGTTTAGAGA 1213
    : : : : : : : : : : : : : : : : : : : : : :
621 tGluLeuTyrGlyLeuValGlnGlnSerHisTyrAlaAspGlnLeuAsp 638
    : : : : : : : : : : : : : : : : : : : : : :
1214 GTTATGCCAGACGCTTTAAAAAATAGCTATGACGAGATATTTATTA 1263
    : : : : : : : : : : : : : : : : : : : : : :
638 yLeuAlaGlnLeuSerSerAlaTyrGlyGlnLysAlaLeuLeu 654
    : : : : : : : : : : : : : : : : : : : : : :
1264 GCAGATATACAGCGGGGACAGGAGTATGATCGTGGTACTGTCATP 1313
    : : : : : : : : : : : : : : : : : : : : : :
655 AlaGlnLeuTyrThrAspLysThrAlaAlaGlnGlyAlaValAlaGly 671
    : : : : : : : : : : : : : : : : : : : : : :
1314 TTAATACCGCATTCGCCGCTATTCGCTGGTGTCT...GCTGCTGAG 1360
    : : : : : : : : : : : : : : : : : : : : : :
671 lSerAlaValLeuSerThrValGlyAlaAlaValSerIleAlaAlaAla 688
    : : : : : : : : : : : : : : : : : : : : : :
1361 CCGATTATACATTGAAAAAGTTAAACATATCTGTCATCGAGATAC 1410
    : : : : : : : : : : : : : : : : : : : : : :
688 lAsSerValValGlyAlaProVal.....AlaValAlaThrSerLeu 701
    : : : : : : : : : : : : : : : : : : : : : :
1411 AAAAAAGACAAAGTGAC.....ATTCAAACACTGCTCCG 1445
    : : : : : : : : : : : : : : : : : : : : : :
702 LeuThrGlyAlaLeuAsnGlyLeuAlaGlyValGlnGlnProIle 718
    : : : : : : : : : : : : : : : : : : : : : :
1446 AGAGGCT.....GATTTGCTAAGACAGTGCCTAAT..... 1476
    : : : : : : : : : : : : : : : : : : : : : :
718 eGluLysLeuAlaAsnAspTyrAlaThrGlySerIleAspGluLeuGly 735
    : : : : : : : : : : : : : : : : : : : : : :
1477 .....TATAAGCACTAATAAGATGAGAAATCGAATAATCATC 1515
    : : : : : : : : : : : : : : : : : : : : : :
735 roGlnAlaTyrPheGluLysAsnLeuGlnAlaThrHisGlnGlnLeuAla 751
    : : : : : : : : : : : : : : : : : : : : : :
1516 GGTCAAAATGGGAGCGGATCACCTCAAGCAAGTGAATGATTTATTCG 1565
    : : : : : : : : : : : : : : : : : : : : : :
752 AsnSerAspGlyLeuArg.....LysMetLeuAlaAspLeuAla 765
    : : : : : : : : : : : : : : : : : : : : : :
1566 AAAAGSTACGGCAAA.....ATTACCAAGATGAGCTATCA 1603
    : : : : : : : : : : : : : : : : : : : : : :
765 acGlyTrpAsnAlaSerSerValIleGlyValGlnThrThrGlnIleSer 782
    : : : : : : : : : : : : : : : : : : : : : :
1604 AAGTGTGTATATCATGTGAATGCTCAAAACATAGCAAAATGAGCAAA 1653
    : : : : : : : : : : : : : : : : : : : : : :
782 ys.....SerAlaLeuGlnLeuAlaAlaIleThrGlyLysAlaAsp 796
    : : : : : : : : : : : : : : : : : : : : : :
1654 AGCTTAGATAGTATATCATCTGATGATGATTT 1689
    : : : : : : : : : : : : : : : : : : : : : :
797 .....LeuLysSerValaAspValPhe 803
    : : : : : : : : : : : : : : : : : : : : : :

```

seq\_name: SwissProt\_39:CYAA\_BORBR  
 seq\_documentation\_block:  
 ID CYAA\_BORBR STANDARD; PRT; 1705 AA.

```

AC 057506; 005179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BIFUNCTIONAL HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (CYCLOLYSIN) (ACT)
DE (AC-HLY) [COMTAINS: CALMODULIN-SENSITIVE ADENYLATE CYCLASE
DE (EC 4.6.1.1) (AMP PYROPHOSPHATE-LYASE) (ADENYLATE CYCLASE); HEMOLYSIN].
CN CYA OR CYAA.
OS Bordetella bronchiseptica.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
RN [1]
RP
RC STRAIN-CIP 9.73;
RC MEDLINE=96009899; PubMed=7557410;
RA Betsou F., Sismelro O., Danchin A., Guiso N.;
RT "Cloning and sequence of the Bordetella bronchiseptica adenylate
RT cyclase-hemolysin-encoding gene: comparison with the Bordetella
RT pertussis gene".
RL Gene 162:165-166(1995).
RN [2]
RP REVISION TO 1517.
RC STRAIN-CIP 9.73;
RA Danchin A.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF
CC BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN
CC CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL
CC CELL FUNCTION.
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.
CC -1- ENZYME REGULATION: ACTIVATED BY HOST CALMODULIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY (BY SIMILARITY).
CC -1- PTM: RELEASED IN A PROCESSED FORM.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO ADENYLATE CYCLASE
CC CLASS-2 FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE RTX
CC PROKARYOTIC TOXIN FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z37112; CAAB5481.1; -
DR INTERPRO; IPR001343; -
DR PFM; PF00353; HemolysinCabin; 5.
DR PRINTS; PR00313; CABDNRGRT.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 5.
KW Lyase; CAMP synthesis; ATP-binding; Hemolysis; Toxin; Virulence;
KW Whooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
FT CHAIN 1 312
FT CHAIN 313 1705
FT FT
FT FT HEMOLYSIN (BY SIMILARITY TO E. COLI
FT FT A, CATALYTIC.
FT FT B, ALA/GLY-RICH.
FT FT C, ASP/GLY-RICH.
FT FT D, ATP (POTENTIAL).
FT FT 28 X REPEATS, GLY-RICH.
FT REPEAT 1014 1019 1.
FT REPEAT 1023 1028 2.
FT REPEAT 1032 1037 3.
FT REPEAT 1041 1046 4.
FT REPEAT 1050 1055 5.
FT REPEAT 1059 1064 6.

```

```

FT REPEAT 1079 1084 7.
FT REPEAT 1164 1169 8.
FT REPEAT 1173 1178 9.
FT REPEAT 1182 1187 10.
FT REPEAT 1202 1207 11.
FT REPEAT 1279 1284 12.
FT REPEAT 1288 1293 13.
FT REPEAT 1297 1302 14.
FT REPEAT 1306 1311 15.
FT REPEAT 1315 1320 16.
FT REPEAT 1324 1329 17.
FT REPEAT 1344 1349 18.
FT REPEAT 1420 1425 19.
FT REPEAT 1429 1434 20.
FT REPEAT 1438 1443 21.
FT REPEAT 1447 1452 22.
FT REPEAT 1455 1460 23.
FT REPEAT 1564 1569 24.
FT REPEAT 1573 1578 25.
FT REPEAT 1582 1587 26.
FT REPEAT 1592 1597 27.
FT REPEAT 1604 1609 28.
FT LIPID 859 864 28.
FT LIPID 982 987 28.
SQ SEQUENCE 1705 AA: 177249 MW: D8530697A4BAE60E CRC64;

```

alignment\_scores:

Quality:	206.50	Length:	372
Ratio:	1.059	Gaps:	11
Percent Similarity:	52.419	Percent Identity:	23.387

alignment\_block:

us-09-306-689-12 x CYAA\_BORBR ..

Align seg 1/1 to: CYAA\_BORBR from: 1 to: 1705

```

300 TTGGAGCTAGCGCTGCGGCTGAGCTTACCTCCAAAACCTGGGG 349
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
380 LeuGluThrValProAlaSer.ProGlyLeuArgArgProSerLeuLys 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
350 CAAAAAATATTCCTATATTCCTCCCAAAATTCACATATGATACGAA 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 LavalGlu.....ArgGlnAspSerGlyTyrAspSerLeu 407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 CAAGGTATGCTTACAGATTAGTCAAGGCGCGAGAGTGGGGAT 449
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 AspGlyValAlaIysSerArgSerPhe..... 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 TGAAGTACAAAGAGAACGCAATTAATTCACACAGCTCAACAGTT 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 ..... 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
500 TAGGACGATTCAAACCGCTATTGGCTTAAGCTAGCGCTGCGATGTGTTA 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 euGlyGluValSerAspMet.....AlaIleVal 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
550 TCGCGCTCCAAATTTGATTAATTCCTACAGAAAACCTAAGAGCGCAAGC 599
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
427 GluAlaIleValLeuGluMetThrArgGlnValLeuHisAlaGlyAlaAr 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 ATTAGGTTCTGCCGAAAGATTTGACAAAATGCAAAATGAA.....GCCA 643
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
443 gGlnAspAlaIleGluProGlyAlaSerGlyAlaSerAlaHisIleTrpGly 460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 AAAGCTATATCTGGCATTCATCTATTTTAGCGTCAGATTGGCTGGA 693
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
460 InArgAlaIleGluGlnGlyAlaIleValAlaIleAlaAla..... 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
694 ATGATTTAGATGAGCGCTTACAGAAATACAGCAACCAATGCTCTTGC 743
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
473 .....GlnArgLeuValHisAlaIleAl 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

744 T.....AAAGCTGCTTGGAGCTAACAAATTCATTAA 775
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
480 aleuMetThrGlnPheGlyArgAlaGlySerThrAsnThrProGlnIleu 497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
776 TTGAAATATTCCTAATTCAGTAAACACTTGACGAAATTTGGGACAA 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
497 laaIleSerLeuSerAlaIleValPheGlyLeuGlyAlaIleSerAla 513
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
826 ATTAGTCAA.....TTGGTTCAAACTACAAATATTCAA 860
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
514 ValAlaIleThrValSerGlyPhePheArgGlySerSerArgTrpAlaGly 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
861 AGCGTTAGGACTTTAGAGACAACTCAAAATATTCGTCGACTTGATA 910
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
530 YGlyPheGlyValAlaGlyGlyAlaMetAlaLeuGlyGlyGlyLeuGly 547
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
911 AAGCTGCGCTTGGTTAGATGTTATCTCAGGCGCTATTATGCGGCGACA 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
547 LavalGlyAlaGlyMetSer..... 553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
961 GCTCAGCTTGTGCTGAGATTAATAATGCTTCAACAGCTAAATAAGTGGG 1010
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
554 .....LeuThrAsp..AspAlaProAlaGlyGlnIleValAla 565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1011 TCGCGCTTTGAAATTCGCAACCAAGTGTGTGTAATTAATTCACAAAGCCG 1060
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
565 aAlaGlyAlaGlnIleAlaLeuGlnLeuThrGlyThrValGlnIleu 582
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1061 TTCTCTTACATTTA.....GCCACAGCTGTTCAGCAGGT 1098
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
582 laSerSerIleAlaLeuAlaLeuAlaIleAlaIleArgGlyValThrSerGly 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1099 TTATCTTACAGTGGCGCTGCTGCTTAAATTCCTAGCTGCTTCTCT 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
599 LeuGlnValAlaGlyAlaSerAlaGlyAlaAlaIleValAlaIleAla 615
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1149 TCGGATTAGCCCATTCATTCATTTGCCGCTATTCGCGATTAATCATG 1198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
615 aAlaLeuSerProMetGlnIleGlyGlyLeuValGlnGlnSerHisIle 632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1199 CAAAAATTTAGAGAGCTTATTCGCAACGCTTAAATAATGCTATGAC 1248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
632 laAspGlnLeuAspLysLeuAlaGlnIleSerAlaIleValGlyGlyGlu 648
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1249 GAGATTAATTTATTCAGCAATATACAGCGGGAACAGGACTATTCATGC 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
649 GlysAlaLeuLeuAlaGlnIleuTyrArgAspLysThrAlaIleGlu 665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1299 ATCGGTTACTGCAATTATACCGCATTTGCCGCTATTCGTTGGTGTGT 1348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
665 yAlaValAlaGlyValSerAlaValLeuSerThrValGlyAlaIleVal 682
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1349 CTGCTGCTGCAGCC 1362
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
682 erIleAlaIleAlaIle 686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq\_name: SwissProt\_39:TALA\_DICDI

seq\_documentation\_block:

```

ID TALA_DICDI STANDARD; PRT; 2491 AA.
AC P54633;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FILOPODIN (TALIN HOMOLOG).
GN TALA.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Dictyostelida; Dictyostelium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2.
RX MEDLINE=95213284; PubMed=7698984;
Kreitelmeier M., Gerisch G., Heizer C., Mueller-Taubenberger A.;

```

"A talin homologue of Dictyostelium rapidly assembles at the leading edge of cells in response to chemoattractant.";  
 Rt Cell Biol. 129:179-188(1995).  
 CC -1- FUNCTION: ACTIN-BINDING PROTEIN THAT MAY BE INVOLVED IN THE  
 CC CONTROL OF CELL MOTILITY AND CHEMOTAXIS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; RAPIDLY ASSEMBLES AT THE  
 CC LEADING EDGE OF CELLS IN RESPONSE TO CHEMOATTRACTANT.  
 CC -1- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,  
 CC RADIXIN, AND TALIN.  
 CC -1- SIMILARITY: IN THE C-TERMINUS TO YEAST SLA2 AND C.ELEGANS ZK370.3.  
 CC PARTIAL; TO YEAST ROD1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U14576; AAC46586.1; -  
 DR DICTYDB: DD01121; TALA.  
 DR INTERPRO: IPR000299; -  
 DR INTERPRO: IPR002558; -  
 DR PFAM: PF00373; Band\_41; 1.  
 DR PFAM: PF01608; L.WEO; 1.  
 DR PROSITE: PS00660; BAND\_41\_1; 1.  
 DR PROSITE: PS00661; BAND\_41\_2; 1.  
 DR PROSITE: PS50057; BAND\_41\_3; 1.  
 DR Structural protein; Cytoskeleton; Actin-binding.  
 KW DOMAIN 134 294 BAND 4.1-LIKE.  
 FT SEQUENCE 2491 AA; 26881 MW; B24FF0224F2A419B CRC64;

alignment\_scores:  
 Quality: 182.50 Length: 650  
 Ratio: 0.583 Gaps: 31  
 Percent Similarity: 48.154 Percent Identity: 22.000

alignment\_block:

US-09-306-689-12 x TALA\_DICDI ..

Align seg 1/1 to: TALA\_DICDI from: 1 to: 2491

```

412 TTACAGAGATTAGTCAAGCGCCGAGAGTTGGGATTGAGTACAAAG 461
    ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1064 leuGlyAspLeuVallySerIatThrGluSerIatSerValAla11 1080
    ::::: ::::: ||::: ||::: ||::: ||::: ||::: ||:::
462 AGAAGAACCAATTAATATGCAACAGCTCAACACGATTAGCGAGATTC 511
    ::::: ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1080 eSerAlaIySthLeuIleCysAla.....ThrtHrGlyLysGlnValG 1095
    ::::: ::::: ||::: ||::: ||::: ||::: ||::: ||:::
512 AAACCGCTATT...GGCTTAAGTACGCTGCGATGTGTATCCGCTCA 558
    ||||| ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1095 LnthrLysLeuMetGlyIleThrLysGlnLeuMetIle..... 1107
    ::::: ::::: ||::: ||::: ||::: ||::: ||::: ||:::
559 CAAATGATGAATATGCTACAGAAACTAAAGCAGCCCAAGCATTAAGTTC 608
    ::::: ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1108 AspMetGlnGlnLeuIleArg.....Al 1115
    ::::: ::::: ||::: ||::: ||::: ||::: ||::: ||:::
609 TGCAGAAACCATTTGTACAAATGAATAAGCCAAACGATTA... 654
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1115 aSerAlaIySerValArgSerAsnProAsnAspArgSerGlnLeu 1132
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
655 .....TCTGGCATTCATTAATTTTGGCTCAGTATTTGGCTGGA 693
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1132 euAspArgArgSerAsnAspValSerIleSerThralaIaleuValGly 1148
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
694 .....ATGGATTAGATGAGCGCTTACAGAAATA 722
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1149 SerThralaAsnValAspCysLysGlnIleAlaSerAlaSerAla 1165
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
723 CAGCAACCAACATGCTCTTGTAAAGCTGGCTTGAGCTAACAAATCA. 771

```

```

1165 eSerAsnLeuLeuSerLeuLysMetGlySerLeuGluSerIleLeuSerG 1182
    ||||| ::::: ||||| |||||
772 .....TAAATGAATAATATGCTAAATTCAGTA 798
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1182 LnthrThrGlnGluPheAlaPheTyValGlnGluIleAlaSerSerThr 1198
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
799 AAACACTTACAGAAATTTGGTACAGCAAAATTAATTCATTTGGTCAAACT 848
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1199 LysAlaLeuAsnAlaIaIaSerGlnGlnValAlaIaMet...AlaArgAs 1214
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
849 ACAAATATCAAGCGCTTAGGACTTAGAGACAAACCAAAATATCG 898
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1214 nlyAsnLeuLysGlnLeuGlyAla..... 1222
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
899 GTGACTTGTAAAGCTGGCTTGTAGATGTATTCAGCGCTATTA 948
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1223 .....SerAlaIySthLeuIleThralaSerAlaLeuSerThrLeuVal 1235
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
949 TCGGGCGCAACAGCTGCATTTGACTT.....GCAGATAAAA 986
    ||| ||| ||||| ||||| |||||
1236 SerHisAlaGlnAsnAlaIleValLeuThrGluAsnGlnLathrLysAs 1252
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
987 TGCTTCACAGCTAAAGAGTGGGCGGCTTTGAAATTCGCAACCAAG 1036
    ||||| ||||| ||||| ||||| |||||
1252 nAlaIleLeuAlaSerThrValAlaLeuGly.....GlyGlnI 1265
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1037 TTTGTGCT..... 1044
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1265 IetleGlyLeuLeuAspPheSerLysAlaArgIleAlaAsnTyrlLysAsp 1281
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1045 .....AATATTACCAAGCGTTTCTCTTA 1070
    ||| ||| ||||| ||||| |||||
1282 ProIleTyAspGlnAsnLeuIleAsnGlnAlaLysSerValGluAspH 1298
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1071 CATTTAGCCCAAGCTGGTGCACGAGTTTAACTCAACGGG..... 1113
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1298 SLeuVal.....LysValGlyArgSerLeuGlyGlyAspGlyAsnAsn 1313
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1114 .....CCTGGCTGCTTAAATGCTTACTGCTTCTCTT 1149
    ||| ||| ||||| ||||| |||||
1313 hrlleCysAspGluAlaValAspArgIleIleGlnLathrArgSerLeu 1329
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1150 .....GCGATTAGCCCATTTGCCGATTTGCCGATAAATTAA 1193
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1330 AspLysThrIleLeuPro...AspThrSerGlyLeuGlnThrAsnAlaH 1345
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1194 T.....CATGCAAAAGTTTAGAGATTATGCCGAGCGCTTA 1231
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1345 SLeuGluMetLeuHisGlnGlnSerLeuLeuAlaIleThrGlnAlaSerL 1362
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1232 AAAAATTTAGGCTATGACGAGATAATTTATAGCAAAATATCAAGCG... 1278
    ||||| ||||| ||||| ||||| |||||
1362 ystLysLeuGlySerIleThrSerAsnLeuValAsnSerLysAsnAsnSer 1378
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1279 .....GGAACAGCGACTATTGATGCA..... 1299
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1379 AspLeuValGlySerGlySerThrAspAlaGluArgIleIleGluMetI 1395
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1300 .....TGCCTTACTGCAATTAATACCG 1321
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1395 eGluAlaIaIaLysHisValAlaHisCysSerIleSerThrTyAsnProA 1412
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1322 CATTTGCCGCTATTTGCTGTGTGCTGTCTGCTGCACCGCATTTAACA 1371
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1412 splleLeuLeuProAlaLysSerIleLeuAspAlaSerGlnMetLeuThr 1428
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1372 TTTGAAAAAGTTAAACATTAATTTGCATCAGCAATAGCAAAAAAGAGA 1421
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1429 AlaAsnGlnAlaAspValAsnHisValLeuSerHisAla..... 1441
    ::::: ||::: ||::: ||::: ||::: ||::: ||:::
1422 AGTGACCATTTCAAAAGCTGGTTCGAGAGCGCTGATTTTGTCAAGAAGTGC 1471

```









CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

-----  
 DR EMBL: M9063; AAA35746.1; -  
 DR INTERPRO: IPR01664; -  
 DR INTERPRO: IPR003054; -  
 DR PFAM: PF00038; filament; 1.  
 DR PRINTS: PR01276; TYPEKERATIN.  
 DR PROSITE: PS00226; IF: 1.  
 KW Intermediate filament; Coiled coil; Heptad repeat pattern; Keratin;  
 KW Phosphorylation.  
 FT DOMAIN 1 182 HEAD.  
 FT DOMAIN 183 492 ROD.  
 FT DOMAIN 493 638 TAIL.  
 FT DOMAIN 183 218 COIL 1A.  
 FT DOMAIN 219 237 LINKER 1.  
 FT DOMAIN 238 329 COIL 1B.  
 FT DOMAIN 330 353 LINKER 12.  
 FT DOMAIN 354 492 COIL 2.  
 FT MOD.RES 57  
 FT SEQUENCE 638 AA: 63871 MW: 987438B872076AF CRC64;  
 alignment\_scores:  
 Quality: 166.00 Length: 734  
 Ratio: 0.505 Gaps: 37  
 Percent Similarity: 44.823 Percent Identity: 23.161

## alignment\_block:

US-09-306-689-12 x K220\_HUMAN  
 Align seg 1/1 to: K220\_HUMAN from: 1 to: 638

```

40 GGCCTCGCCCTGGCAGCGTCTCTCAAGATTGGAGCTACGCGCTGGT... 87
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
45 GlyPheArgSerCylValagly.....SerPheGlySerArgse 57
88 .....CCGGTG 94
57 rleuTyraAsnLeuGlySerAsnLysSerIleSerIleSerValAlaIag 74
95 GGTCTAGCCAGCATTTGGAGCTACGCGCTCGCCCTGACGCGGTACCA 144
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
74 LysSerSerArgAlaGlyGlyPheGlyGlyArgSerSerCylsGlyPhe 90
145 GATTGGAGCTACGCGCTCGCGCTCGGTGATCTCAGCATTTGAGCTACG 194
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
91 AlaGlyGlyTyrglyGlyGlyPheGlyGlySerTyrglyGlyGlyPheG 107
195 CCTGGCCCTGCGCAGCGCTTCTCAAGATTGAGCTACGCGCTCGCG 244
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
107 yGlyGlyArgGlyValGlySerGlyPheGlyGlyAlaGly.....GlyPhe 123
245 GTGGCTCTAGCCAGCATTTGGAGCTACGCGCTCGCCCTGACGCGGTAC 294
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
123 yGlyValAlaGlyGlyPheGlyGlyProGlyValPheGlyGlyProGlySer 139
295 CAAGATTGGAGCTACGCGCTCGCGCTCGGTGATCTCAGCATTTCCCA 344
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
140 PheGlyGlyProGlyGlyPheGlyProGlyGly.....PhePro...G 153
345 TGGGGCAAAAAAATTATC.....CTGATATTTCCCAAAAT 382
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
153 yGlyIleGlnGlnValIleValAsnGlnSerLeuGlnProLeuAsnV 170
383 ACCAATATGATCTGAAACAAGTAATGTTACAGATTTCAGCAAGG 432
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
170 alGluIleAsnProGlnIleGly..... 177
433 GCCGAAGAGCTGGGATGGAGTACAAAGAGAAGCAATATATATTCG 482
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

178 .....GluValLysAlaGlnGluArgGlnIleGly 188
483 AACAGCTCAACACGATTAGCAG...ATTCAACCGCTATTGGCTTAA 529
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
188 sTrIleuAsnAsnLysPheAlaSerPheIleAsnLysValArgPheGlu 205
530 CTGAGCGCTGATTTGTTATCGCGCTCCACAAATTCATTAATTCGTAC 579
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
205 LuGlnGlnAsnLysValLeuGlnThrLysTrp.....GluLeuGln 219
580 AAACCTAAACAGCCAGCATTTAGTCTGCC.....GAAG 617
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
220 GlnGlnThrThrLysSerGlyProSerSerLeuGlnProCysPheGlu 236
618 CATTTGACAAATATGCAATTAAGCCAAACCTGATTATTCGTGCAATCA 667
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
236 rTyrlleSerPheLeuGlySerGlnLeuAsnSerLeuGlyGluArg 253
668 CTATTTTAGGCTCAGTATTCGCTGCAATG.....GATTA 702
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
253 LysAsnLeuGlnGlyGluLeuLysSerMetGlnAsnLeuValGluAsn 269
703 GATGAGCGCTTACAGATAACACCAACCAACATGCTCTTCTTAA..... 747
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
270 LysLysLysTyrglyLysAsnLysValArgThrAlaAlaGlnGln 286
748 ...GCTGCTTGAGCTTAACAAATTCATTAATTAATTAATTCATTA 793
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
286 upheValGlyLeuLys..... 291
794 CAGTAATAAACACTTGACGAATTTGGTACCAATTTGCTAATTTGCTTCA 843
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
292 ....LysAspValAsp.....AlaAlaPheMetAsn 300
844 AAACCTCAAAATATCAAGGCTTACGAGCTTACGACCAACTCAAAAA 893
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
301 LysValGlnLeuGlnAlaLysValAspSerLeuThrAsnGlnValSer 317
894 TATCGGTGACCTGATTAAGCTGCGCTTGATTGATTATCTCAGGCG 943
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
317 eleuArgThrLeuTyrglyMetGlnLeu.....SerGln 329
944 TATTATCGGCGCA.....ACAGCTGCACTGTACTTCAGATFAAAAT 987
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
329 eteInserHisAlaSerAspThrSerValValLeuSerMetAsnAsn 345
988 GCTTCAACAGCTPAAAAAGTGGTGGGCTTTGAAATTCGCAACCAAG 1037
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
346 ArgCysLeuAsp..... 349
1038 TGTGGTAATTAATTAACCAAGCGTTTCTTCTTACATTT.....TTAGCCC 1081
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
350 LeuGlySerIleIleAlaGlnValArgThrGlnTyrglnGlnIleAla 366
1082 AACGTTTGACAGAGTTATCTTCAACTGCGGCTTGCTGCTTTAAT 1131
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
366 lnaTy.....SerLysSerGlnAlaGlnAlaLeuTy 376
1132 GCTTCTACTGTTCTCTTCGATTAACCATTAAGCATTTTGGCGTATTCG 1181
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
377 GlnThrLysLeuGly..... 381
1182 CGATAAATTAATCATGCAAAAAAGTTTAGAGATTATGCGGAACGCTT 1231
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
382 .....GluLeuGlnThrAla..... 387
1232 AAAAATTAAGCTATGACGAGATTAATTTATAGCAGAAATATACGCGG 1281
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
388 .....GlyArgHisGlyAspAspLeu.....ArgAsn 396
1282 ACAGGACTATGATGATCGGTACTGCAATTAATACGCAATTCGCGCCG 1331
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
397 Thr.....LysSerGlnIleMetGlnLeuAsnArgMetIleGln 410

```

```

1332 TATTGCTGCTGCTGCTGCTGCTGCGAGCCGATTAACTTGAAG 1381
410 gldu..... 411
1382 TTAACATATCTTGTATCAGCAATACCAAAAAAGAGAGTCCATT 1431
412 .....ATGAlaGluIleGluIleGluIleGluIleGluIleGluIle 425
1432 CAAAGCTGCTGCGAGCGCTGATTGCTTAAAGAGTCCATTATTA 1481
426 GlnThrAlaIleAlaGluAlaGluIleGluIleGluIleGluIle 438
1482 AGCACTAAAGATGAGAAATCGAAGAAATCATGCTCAAAATGCGAGC 1531
439 AlaLeuLysasp..... 442
1532 GGATCACCCTCAAGCAAGTGTATCTT.....ATGCAAAAGCT 1572
443 .....AlaAsnAlaLysLeuGlnAspLeuGlnThrAlaLeuGlnLys... 456
1573 AAGGCAAAATATACCAAGATGCTATCAAAAGTGTGTAATCTATGA 1622
457 .....AlaLysAspAspLeuAlaArgLeuLeuAlaArgAspTyrGlu 469
1623 ATTGCTCAAAATAGCAAAATGTGACAAACAGCTTAGAT..... 1662
469 ngldu.....MetAsnValLysLeuAlaLeuAspAlaGluIleA 483
1663 .....ACGTATCTCATCTGTAGTGCATTTACTCTGCTCAT 1701
483 lathrTyrArgLysLeuLeuGluGluGluGluGluGluGluGlu 499
1702 GATTCGAGAAATGATTAGTGTCTCAACTCAATGTTGATCAAGT 1751
500 CysIleSerAlaValLysIleSerValValSer.....AsnVal 512
1752 ATCTTCTCTTCAATTTCTAGGAGATCTCAGCATTTGAGACGCGCTGC 1801
512 lThSerThrSerLysSerLysSerLysSerLysSerLysSerLys 529
1802 GCCCTGCGAGCGCTTCTCAAGATGAGCTACGCGCTCGCGGTCGC 1851
529 alSerGlySerGlySer.....GlyGlyTyrLysGlyGly 540
1852 TCTAGCCGAGCTTGGAGC.....TAGGCGCTGCGCGCTGCGAGCGGTG 1895
541 SerSerSerSerSerSerGlyTyrGlyValSerGlyLysSerGly 557
1896 CCAAGATTGAGCTACGCGCTGCTCGCGGATCTCAGCATTTGAGCT 1945
557 r.....GlyTyr.....GlyGlyValSerSerGlySerT 567
1946 ACGGCTGCGCGCTGCGAGCGCTTCTCAAGATTGAGCTACGCGCTG 1992
567 hrgLysArgLysArgLysSerSerGlySerSerSerSerSerSer 583
1993 CGTCGCGGTGC.....TCTAGCCGAGCTTGGAGCTACGCGCT 2030
584 ArgLeuGlyAlaGlySerIleSerValSerIleSerGlyMetGly 600
2031 GCGGCTGCGAGCGCTGAGCCAAAGATTGAGCTACGCGCTGCGGTG 2080
600 rSerSerGlySerIleGlnThrSerGlyLysSerGlyTyrLysSerGly 617
2081 GA 2082
617 ly 617
seq_name: SwissProt_39:K22E_HUMAN
seq_documentation_block:
ID K22E_HUMAN STANDARD: PRT; 645 AA.

```

```

AC P35908;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KERATIN, TYPE II CYTOSKELETAL 2 EPIDERMAL (CYTOKERATIN 2E) (K2E)
DE (CK 2E).
GN KRT2A OR KRT2E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HIGH EPIDERMIS;
RX MEDLINE=92380238; PubMed=1380918;
RA Collin C., Moll R., Rubicka S., Ouhayoun J.-P., Franke W.W.;
RT "Characterization of human cytokeratin 2, an epidermal cytoskeletal
RT protein synthesized late during differentiation."
RL Exp. Cell Res. 202:132-141(1992).
RN [2]
RP SEQUENCE FROM N.A. AND VARIANTS IBS TYR-192 AND LYS-482.
RA Smith F.J.D., Maini C., Coveello S.P., Higgins C., Schmidt M.,
RA Lane E.B., Uitto J., Leigh I.M., McLean W.H.I.;
RT "Genomic organization and fine mapping of the keratin 2e gene (KRT2E):
RT of Siemens."
RL J. Invest. Dermatol. 111:817-821(1998).
RN [3]
RP VARIANT IBS LYS-493.
RX MEDLINE=94358459; PubMed=7521371;
RA McLean W.H.I., Morley S.M., Lane E.B., Eady R.A., Griffiths W.A.,
RA Pidge D.G., Harper J.I., Higgins C., Leigh I.M.;
RT "Ichthyosis bullosa of Siemens -- a disease involving keratin 2e."
RL J. Invest. Dermatol. 103:277-281(1994).
RN [4]
RP VARIANTS IBS PRO-187 AND LYS-493.
RX MEDLINE=94358461; PubMed=8077693;
RA Kremer H., Zeeuwen P., McLean W.H.I., Mariman E.C.M., Lane E.B.,
RA van de Kerkhof P.C.M., Ropers H.-H., Steijlen P.M.;
RT "Ichthyosis bullosa of Siemens is caused by mutations in the keratin
RT 2e gene."
RL J. Invest. Dermatol. 103:286-289(1994).
RN [5]
RP VARIANTS IBS ASP-493 AND LYS-493.
RX MEDLINE=95038833; PubMed=7524919;
RA Rothnagel J.A., Traupe H., Wojcik S., Huber M., Hohl D.,
RA Pittelkow M.R., Saeki H., Ishibashi Y., Roop D.R.;
RT "Mutations in the rod domain of keratin 2e in patients with ichthyosis
RT bullosa of Siemens."
RL Nat. Genet. 7:485-490(1994).
RN [6]
RP VARIANT IBS PRO-485.
RA Yang J.-M., Lee S., Bang H.-D., Kim W.-S., Lee E.-S., Steinert P.M.;
RT "A novel threonine-to-proline mutation at the end of 2b rod domain in
RT the keratin 2e chain in ichthyosis bullosa of Siemens."
RL J. Invest. Dermatol. 109:116-118(1997).
CC -1- FUNCTION: PROBABLY CONTRIBUTES TO TERMINAL CORNIFICATION.
CC -1- SUBUNIT: HETEROHEXAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- TISSUE SPECIFICITY: IN THE SUPRABASAL LAYERS OF EPIDERMAL
CC TISSUES FROM MOST BODY SITES, EXCEPT IN FORESKIN, IN SQUAMOUS
CC METAPLASIAS AND CARCINOMAS.
CC -1- DEVELOPMENTAL STAGE: SYNTHESIZED DURING MATURATION OF EPIDERMAL
CC KERATINOCYTES.
CC -1- DISEASE: DEFECTS IN KERATIN 2E ARE ASSOCIATED WITH ICHTHYOSIS
CC BULLOSA OF SIEMENS, A RARE AUTOSOMAL DOMINANT DISORDER DISPLAYING
CC A TYPE OF EPIDERMOLYTIC HYPERKERATOSIS CHARACTERIZED BY EXTENSIVE
CC BLISTERING FROM BIRTH. HYPERKERATOSIS AND SHEDDING OF THE OUTER
CC LAYERS OF THE EPIDERMIS (MOLTING) ARE OBSERVED IN LATER WEEKS.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFIBRILLAR KERATIN: I (ACIDIC: 40-55 KDA) [K9 TO K20] AND II
CC (NEUTRAL TO BASIC: 56-70 KDA) [K1 TO K8].
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



```

368  rgseryslglulalaglualeuylrhisserlys..... 380
1216  TATGCCAAGCCTTAAATAATAGCTATGACGAGATATTTATAGC 1265
381  TTTTCTGluLeuGluValThrValGlyArgHisGlyAspSerLeu..... 395
1266  AGAATATCAGCGGGAACAGGAGCTATTGATGCATTCGTTACTGCAATTA 1315
396  .....lysGluIleIysIleGluIleSerLeuLeuA 406
1316  ATACCGCATGGCGGCTATGCTGGTGTGTCTGCTGTCAGCCGAT 1365
406  snarValIleGluArgLeuGlnGly..... 414
1366  TTAACATTGAAAAGTTAAACATATCTGTCTACGATACCAAAA 1415
415  .....GluIleAlaHisValLysLys 421
1416  AGAGAAATGACCATTCAAACCTGTCGAGAGCTGATTTCTTAAG 1465
421  sGluCylLysAsnValGlnAspAlaIleAlaAspAlaGluArgGlyG 438
1466  AAGTGCCTAATTATAAGCACTAAAGATGAGAAA..... 1500
438  Lu.....HisAlaLeuLysAspAlaArgAsnLysLeuAsnAsp 450
1501  ATCAAGAAATCATCGGCAAAATGGCGAGCGGATCACCTCAAGCAAGT 1550
451  LeuGluIleAlaLeuGlnGln..... 457
1551  TGATGATCTTATCGCAAAAGTAAACGCAAAATTAACCAAGATGAGTAT 1600
458  .....AlaLysLysLeuAsn 463
1601  CAAAAGTTGTGATTAATGATGATGCTCAAAACATAGCAAAATGTGACA 1650
463  laArgLeuLeuArgAspTyrGlnLeu.....MetAsnValLys 476
1651  AACGCTTAGT.....AAGTTAATCTCATCTCT 1679
477  LeuAlaLeuAspValGluIleAlaThrTyrArgLysLeuGlnGlyG 493
1680  AAGTCAATTTACTGCTCTAATGATGATGAGAATATTAGTGGCTCAA 1729
493  uGluCylArgMetSerGlyAspLeuSerSerAsnValThrAlaSerValT 510
1730  CTTCATGTTGGATCAAGATTATCTCTTCAATTGCTAGGGGATCT 1779
510  hrSer.....SerThrIleSerSerAsnValAlaSerLysAla... 522
1780  CAGATTGAGACTACGGCTCGCCGCGGCGGCTTCTCAAGATGTGAG 1829
523  .....AlaPheGlyLysSerGlyLysArgLysSer.....Se 534
1830  CTACGCGCTGCTCGGCTGCTAGCAGCATTTGAGTACGAGCCGTC 1879
534  rGlyGlyLysTyrSerSerGlySerSerTyrGlySerGlyLysArgG 551
1880  GCCCTGAGCGGTAGCCAGATGAGCTACGCGCTGCGGAGTGA 1929
551  InSerGlySerArgGlyLysSerGlyGlyGlySerIleSerGlyLys 567
1930  TCTGAGATTGAGCTACGCGCTGCGGCTGCTTCTCAAGATG 1979
568  .....GlyTyrGlySerGlyLysArgLysSerGlyLysArgTyrG 580
1980  GAGCTAGCGCTGCGGCTGCGGCTGCT..... 2007
580  ySerGlyLysSerLysSerLysSerIleSerGlyLysTyrGlyLys 597
2008  .....AGCCAGCATTTGAGCTACGCGCTGCGGCTGCGGAGCGGTA 2049

```

```

597  erglyGlyLysHisSerSerGlyLysArgGlyLysSer 613
2050  CAAGATTGAGCTACGCGCTGCGGCTGAGTCC 2085
614  SerGlyLysTyrGlySerGlyLysSer 625
seq_name: SwissProt_39:K1C1_HUMAN
seq_documentation_block:
ID      K1C1_HUMAN      STANDARD:      PRT:      622 AA.
AC      P35527, Q14665;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DE      15-JUL-1998 (Rel. 36, Last annotation update)
DE      KERATIN, TYPE I CYTOSKELETAL 9 (CYTOKERATIN 9) (K9) (CK 9).
GN      KRT9.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC      TISSUE=FOOT SOLE TISSUE;
RX      MEDLINE=94131202; PubMed=7507869;
RA      Langbein L., Heid H.W., Moll I., Franke W.W.;
RT      "Molecular characterization of the body site-specific human epidermal
RT      cytokeratin 9: cDNA cloning, amino acid sequence, and tissue
RT      specificity of gene expression.";
RL      Differentiation 55:57-72(1993).
RN      [2]
RP      SEQUENCE OF 449-465.
RX      MEDLINE=90267446; PubMed=2140676;
RA      Rosen E.M., Meromsky L., Romero R., Selter E., Goldberg I.;
RT      "Human placenta contains an epithelial scatter protein.";
RL      Biochem. Biophys. Res. Commun. 168:1082-1086(1990).
RN      [3]
RP      VARIANTS EPPK LYS-160; GLN-162 AND TRP-162.
RX      MEDLINE=94214498; PubMed=7512862;
RA      Reis A., Hennies H.-C., Langbein L., Digweed M., Mischke D.,
RA      Dreschler M., Schroek E., Royer-Pokora B., Franke W.W., Sperling K.,
RA      Kuester W.;
RT      "Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma
RT      (EPPK).";
RL      Nat. Genet. 6:174-179(1994).
RN      [4]
RP      VARIANTS EPPK TRP-162 AND SER-167.
RX      MEDLINE=95164983; PubMed=7532199;
RA      Rothnagel J.A., Wojcik S., Diefer K.M., Dominey A.M., Huber M.,
RA      Hohl D., Roop D.R.;
RT      "Mutations in the 1A domain of keratin 9 in patients with
RT      epidermolytic palmoplantar keratoderma.";
RL      J. Invest. Dermatol. 104:430-433(1995).
CC      -1- FUNCTION: MAY SERVE AN IMPORTANT SPECIAL FUNCTION EITHER IN THE
CC      MATURE PALMAR AND PLANTAR SKIN TISSUE OR IN THE MORPHOGENIC
CC      PROGRAM OF THE FORMATION OF THESE TISSUE.
CC      -1- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC      -1- TISSUE SPECIFICITY: EXPRESSED IN THE TERMINALLY DIFFERENTIATED
CC      EPIDERMIS OF PALMS AND SOLES.
CC      -1- DISEASE: DEFECTS IN KRT9 ARE A CAUSE OF EPIDERMOLYTIC PALMOPANTAR
CC      KERATODERMA (EPPK), AN AUTOSOMAL DOMINANT DISEASE CHARACTERIZED BY
CC      DIFFUSE THICKENING OF THE EPIDERMIS ON THE ENTIRE SURFACE OF
CC      PALMS AND SOLES SHARPLY BORDERED WITH ERYTHEMATOUS MARGINS.
CC      -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC      MICROFILAMENTAR KERATIN: I (ACIDIC; 40-55 kDa) [K9 TO K20] AND II
CC      (NEUTRAL TO BASIC; 56-70 kDa) [K1 TO K8].
CC      -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC      -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE A 60 KDA CHAIN OF
CC      PLACENTAL SCATTER PROTEIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR EMBL: X75015; CAA52924.1; -.  
DR EMBL: Z29074; CAA82315.1; -.  
DR EMBL: S69510; AAC60619.1; -.  
DR PIR: B35494; B35494.  
DR HSPD: P02876; 2MGC.  
DR MIM: 144200; -.  
DR INTERPRO: IPR001664; -.  
DR INTERPRO: IPR002957; -.  
DR Pfam: PF00038; filament; 1.  
DR PRINTS: PR01248; TYPE1KERATIN.  
DR PROSITE: PS00226; IF; 1.  
KW Intermediate filament; Coiled coil; Heptad repeat pattern; Keratin.  
KW Disease mutation.

FT	DOMAIN	1	151	HEAD.
FT	DOMAIN	152	460	ROD.
FT	DOMAIN	461	622	TAIL.
FT	DOMAIN	152	187	COTL 1A.
FT	DOMAIN	188	206	LINKER 1.
FT	DOMAIN	207	298	COTL 1B.
FT	DOMAIN	299	321	LINKER 12.
FT	DOMAIN	322	460	COTL 2.
FT	DOMAIN	14	20	POLY-GLY
FT	VARIANT	160	160	N -> K (IN EPPK).
FT	VARIANT	162	162	/FTId-VAR_003822.
FT	VARIANT	162	162	R -> Q (IN EPPK).
FT	VARIANT	162	162	/FTId-VAR_003823.
FT	VARIANT	167	167	R -> W (IN EPPK).
FT	VARIANT	167	167	/FTId-VAR_003824.
FT	VARIANT	167	167	L -> S (IN EPPK).
FT	VARIANT	167	167	/FTId-VAR_003825.
FT	CONFLICT	12	12	T -> SR (IN CAA82315).
QO	SEQUENCE	622 AA;	61987 MM;	898C3825DA85ED94 CR664;

```
alignment_scores:
  Quality: 163.00      Length: 7411
  Ratio: 0.487        Gaps: 33
Percent Similarity: 45.209  Percent Identity: 20.648
```

```

alignment_block:
US-09-306-689-12 x KICL_HUMAN      ..
Align seg 1/1 to: KICL_HUMAN from: 1 to: 622

```

```

22 TCTCAGCAATTGAGGCTACGGCCCTGCGCCAGGCT...TCTCA 65
   |||||  ||  |||  |||  |||  |||  |||  |||
32 SerArgPheSerSerSerGlyArgGlyGlyGlyArgPheSer 48
   |||||  |||||  |||||  |||||  |||||  |||||
66 AGATTGAGCTACGGCCCTGCGGTGCTCAGCAGCATGGAGCT 115
   |||||  |||||  |||||  |||||  |||||  |||||
78 rSerSerGlyArgGly.....GlyGlySerSerArgValCysGly 62
   |||||  |||||  |||||  |||||  |||||  |||||
116 ACGGCTCGGCCCTGGCAGCGGTAGCCAAAGTTGGACTACGGCTCGCT 165
   |||||  |||||  |||||  |||||  |||||  |||||
62 rgGly.....GlyGlyGlySerPheGlyTrSerTrGlyGlyGly 75
   |||||  |||||  |||||  |||||  |||||  |||||
166 CCGGCTGATCT...CAGCATTTGAGCTACGGCCCTGCGCCAGCAGG 212
   |||||  |||||  |||||  |||||  |||||  |||||
76 SerGlyGlyGlyPheSerAlaSerSerLeuGlyGlyGlyPheGlyGly 92
   |||||  |||||  |||||  |||||  |||||  |||||
213 TTCTCAATTTGAGCTACGGCCCTGCGGTGCTCAGCAGCATT 262
   |||||  |||||  |||||  |||||  |||||  |||||
92 ySerTrGlyGlyPheGly...GlyAlaSerGlyGlyGlyTrSerSerG 108
   |||||  |||||  |||||  |||||  |||||  |||||
263 GGAGCTACGGCCCTGCGCCCTGCGAGCGGTAGCCAAAGTTGGAGCTACGCG 312
   |||||  |||||  |||||  |||||  |||||  |||||
108 LyGlyPheGlyGlyGlyPheGlyGlySer.....GlyGlyGly 121
   |||||  |||||  |||||  |||||  |||||  |||||
313 CTGCGCTCGGGT...GGATTAGCTTCCCAAAAACGTGGGCAAAAAAAT 355
   |||||  |||||  |||||  |||||  |||||  |||||

```

122 phegiylgiyltyrlygyserylphenylglyleuylgly. .... 135

360 TATCCTATATATCCCAAAATTAACCATATGATACGTGAAGSTAATG 409  
::: |||:::|

136 ..... PheGlyIyGlyIaGlyIyG 143

410 GTTTACAGATTTAAGTCAAAAGCGCCGAAGATTGGGATTGAGTACAA 459  
||| ::::: ||| |||:::|

143 IAspIyGlyIleLeuThrIlaaGnIuLyS.....SerthMet 156

460 AAGAAGAAGCAATTAATTTTCAACAGCTCAAAACGATTAGGCAGAT 509  
::: ||| ::::: |||:::|

157 GInGIuLeuAsnSerArgIeuAlaSer ..... 165

510 TCAAAACCGCTATTGGCTTAACGTAGCCGTGATGTATCCGCTCAC 559

166 ..... T 166

560 AAATGATTAATTTGCTACAGAAACTAAAGCGCAAGCATTAAGTCT 609  
::: |||:::|

166 yIleuAspIyVal.....GlnAlaIeuIuG 175

610 GCCGAAGCATTTGCAAAATGCAATATAAGCCAAACGTATTATTCG 659  
||| AAsnAsn..... 178

660 CATTCATATTTATTTAGGCTCAGATTGGCTGAATGATTAGATGAG 709  
::: |||:::|

179 ..... AspIeuGluAsnL 183

710 CCTTACAGATTAACAGCAACCAACATGCTTTGCTAAAGCTGGCTGAG 759  
::: |||:::|

183 yAlIeGInAspTrpIyAspIySgIyProIa..... 194

760 CTAACAATTCATTATTAAGAAATTTGCTATTCAGTAAACACCTTGA 809  
::: |||:::|

195 ..... AlaIeGInIuAsnIySerProIyTrpAsnThrIleAs 208

810 CGAATTTGGTGAACAATTAAGCAATTTGGTTCAAAACTACAAATATCA 859  
|||:::|

208 pAspIeuAspArgInIleValAspLeu..... 217

860 AAGGCTTAGGCACTTAGAGACAACACTCAAAATATCGGTGACTGAT 909  
|||:::|

218 .....ThrValGIyAsnAsn...LysThrIleuAspIleAsp 229

910 AAAGCTGGCTTGTTAGATGTTATTCACAGGGCATTAATCGGGCGCAAC 959  
::: |||:::|

230 AsnThrArgMetThrIleuAspAspPheArg..... 239

960 AGCTGCACCTTGACTTGCAGATTAATAATGCTCAACAGCTAATAAGTGG 1009  
::: ::::: |||

240 .....IleIySpheGInMetGInIaAsnIeuArgGInGlyAla 253

1010 GTGCGGGTTTGAATTTGCAACCAAGTTGTGTATATATACCAAAAGCC 1059  
||| ::::: |||:::|

253 sPAlaAspIleAsnGlyLeuArgGInValIleuAspAsnLeuThr..... 267

1060 GTTTCTCTTACATTTTAGCCCAAGCTGTGCAGCAGGTTATATCTTCAAC 1109

267 ..... 267

1110 TGGGCTGTGGCTGCTTAATGTTCTTACAGTTTCTGTGGATTAAGCC 1159

268 .....MetGIuLySerAspIleuIuMetGInIyGInuT 279

1160 CATTAAGCATTTGCGCGATTTGCCGATTAATTAATCATGCAAAAGTTTA 1209  
||| ::::: ||| |||:::|

279 hIleuGInIuGInIuIeuMetAlaIeuLySAsnIhIySleuGInuMet 295

1210 GAGAGTTATGCCGAAGCGCTTTAAAAAATTAAGGATATGACGGAGATATT 1255  
::: ::::: |||:::|

296 SerGInuLeuThr.....GIyGInAsn... 302

```

1260 ATTAGAGAAATATCAGCGGGAACAGACACTATGATGATCGGTACTG 1309
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
303 .....SerGlyAspValAsnValGlnLeuSrv 312
1310 CAATTAAATACCGCATGCGCGCTATTGCTGCTGCTGCTGCTGCA 1359
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
312 a1.....AlaProGly 315
1360 GCCGATTATTAACATTGAAAAAGTTAAACATTAATCTGTCATCAGAAATG 1409
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
316 LysAspLeuThr.....LysThrLeuAsnAspMet.....ArgG1 327
1410 CAAAAGAGAGAAAGATGACCAATTCAAATCGTGTCGAGAGCGGTGATTTTG 1459
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
327 ngLutylGlnGlnLeuLeuLeuAlaLysAsn.....A 337
1460 CTAAAGAGAGTCCCTAATTATTAACCACTAAAGATGAGAAATGCAAGAA 1509
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
337 rglYsAspIleGlnAsnGlnTyrGlnThrGlnIleThrGlnIleGlnHis 353
1510 ATCATCGGTCAAAATGGCGAGCGGATCACCCTCA.....AAGCAAGTTGA 1553
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
354 GluValSerSerSerGlnGlnValGlnSerSerAlaLysGlnValTh 370
1554 TGATCTATGCGCAAAAGTAAAGCGCAAAAT.....ACCCAGATG 1594
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
370 rGlnLeu...ArgHisGlnValGlnGlnLeuGlnIleGlnLeuLeuSerg 386
1595 AGCTATCAAAAGTTGTTGATTAATACTGAATTCGCAAAATGCAAAAT 1644
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
386 InLeuSerLysLysAlaLeuGlnLysSerLeuGlnAspThrLysAsn 402
1645 .....GTGACAACAGCTGATATGATTAATCTGATCTGATG 1682
403 ArgTyrGlyGlnGlnLeuGlnMetIleGlnGlnIleSerAsnLeuG1 419
1683 TGCA.....TTTACCTGCTGATGATGATGCA 1708
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
419 uAlaGlnIleThrAspValArgGlnGlnIleGlnGlnAsnGlnGln 436
1709 GAAATGATTATGTCGCTCCAACTTCAATGTTGATCAAAAGTTATCTTCT 1738
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
436 YrSerLeuLeuSerIleLysMetArgLeuGlnLysGlnIleGlnThr 452
1759 CTGTAA..... 1784
453 TyrHisAsnLeuLeuGlnGlnGlnAspPheGlnSerSerGlyAl 469
1765 .....TTTGGTAGGGATCTCAGCATTTGAGTACGAGCC 1798
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
469 aglyLysIleGlnLeuGlnGlnArgGlnGlnSerGlyLysSerGly 485
1799 TGGCCCTGCGAGCGGTTCTCAGATGAGCTACGCGCTCGCGGCT 1848
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
486 .ArgGlySerArgGlyLysSerGlyLysSerGlyGlnGlnGlnGln 501
1849 GGCCTAGCCCGCATGAGCTACGCGCTCGCGCTCGCGAGCGAGTACCA 1898
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
502 GlnGlyLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 517
1899 AGATTGAGCTACGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 1948
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
518 .SerTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 534
1949 GCGTGGCCCTGCGAGCGGTTCTCAGATGAGCTACGCGCTCGCGCTCG 1998
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
534 GlnGlyLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 550
1999 GGTGGCTCTAGCCCGCATGCG..... 2019
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
551 GlnGlyLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 567

```

```

2020 .....AGCTACGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 2062
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
567 rglYlGlyGlnTyrGln.....GlnGlySerGlnGlnGlnGlnGlnGln 582
2063 ACGGCTGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 2085
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
582 LysGlnSerHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln
seq_name: SwissProt_39:GUNB_CELFI
seq_documentation_block:
ID GUNB_CELFI STANDARD: PRT: 1045 AA.
AC P26225;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ENDOGLUCANASE B PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE B)
DE (CELLULOSE B).
GN CENB.
OS Cellulomonas fimi.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococccaceae; Cellulomonadaceae; Cellulomonas.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-91100298; PubMed-1987122;
RA Melnik A., Braun C., Gilkes N.R., Kilburn D.G., Miller R.C. Jr.,
RA Warren R.A.J.;
RT "Unusual sequence organization in Cemb, an inverting endoglucanase
RT from Cellulomonas fimi.",
RL J. Bacteriol. 173:308-314(1991).
RN (2)
RP DOMAINS.
RX MEDLINE-92041609; PubMed-1938913;
RA Melnik A., Gilkes N.R., Kilburn D.G., Miller R.C. Jr., Warren R.A.J.;
RT "Multiple domains in endoglucanase B (Cemb) from Cellulomonas fimi:
RT functions and relatedness to domains in other polypeptides.",
RL J. Bacteriol. 173:7126-7135(1991).
CC -1- FUNCTION: THE BIOLOGICAL CONVERSION OF CELLULOSE TO GLUCOSE
CC GENERALLY REQUIRES THREE TYPES OF HYDROLYTIC ENZYMES:
CC (1) ENDOGLUCANASES WHICH CUT INTERNAL BETA-1,4-GLUCOSIDIC BONDS;
CC (2) EXOCELLULOBRONOLASES THAT CUT THE CELLULOSE POLYMER CHAIN;
CC FROM THE NONREDUCING END OF THE CELLULOSE POLYMER CHAIN;
CC (3) BETA-1,4-GLUCOSIDASES WHICH HYDROLYZE THE CELLULOSE AND OTHER
CC SHORT CELLO-OLIGOSACCHARIDES TO GLUCOSE.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC -1- MISCELLANEOUS: THE LINKER REGION (ALSO TERMED "HINGE") MAY BE A
CC POTENTIAL SITE FOR PROTEOLYSIS.
CC -1- MISCELLANEOUS: MAY CONTAIN A SECOND CBD IN THE CATALYTIC DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY E (FAMILY 9 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M64644; AAA23086.1; -
CC PIR: A39199; A39199..
CC HSSP: P26221; 3TF4.
CC INTERPRO: IPR001701; -
CC INTERPRO: IPR001777; -
CC INTERPRO: IPR001919; -
CC PFM: PFO0553; CBD_2; 1.
CC PFM: PFO0759; Glyco_hydro_9; 1.
CC PFM: PFO0041; fn3; 3.
CC PROSITE: PS00561; CBD_BACTERIAL; 1.
CC PROSITE: PS00592; GLYCOSYL_HYDROL_F9_1; 1.
CC PROSITE: PS00698; GLYCOSYL_HYDROL_F9_2; 1.
CC PROSITE: PS00698; GLYCOSYL_HYDROL_F9_2; 1.

```





```

775 AspProAlaValSer..... 779
1194 TCATGCAAAAGTTAGAGAGTTATGCCGACGCTTAAATAATAGGCT 1243
780 .....GLYT 781
1244 ATGACGCGATTAATTATATAGCAAAATATACGCGGAGACGAGGACTATT 1293
781 yrAsp.....ValLeuArgValGlnGlyThrThrThrThrVal 793
1294 GATGCAATCGGTTACTGCAATTAATACCGCATGCGCGCT..... 1332
794 ValAlaGlnThrThrValProThrValThrLeuSerGlyLeuThrProse 810
1333 .....ATGCGTGGCGTGTCT 1348
810 rThrAlaTyrThrTyrAlaValArgAlaLysAsnValAlaGlyAspValS 827
1349 CTGCTGCTGCGCGGATTTAATGAAATTTGAAAGTTAAACATTAATCTTGT 1398
827 erAlaLeuSerAlaProValThrPhe..... 835
1399 ATCAGCAATGCAAAAGAGAAAGTGCATTCACAAATGCTTCGAGAGA 1448
836 ...ThrThrAlaAlaProProValAspThrVal..... 845
1449 GCGGTATTTTGCTAAAGAGAGTGCCTATTTAAAGCACTAAAGATGAGA 1498
846 .....AlaProThrValProGlyThrProValAlaSerAsn.... 857
1499 AATATGCAAGAAATCATGCTCAAAATGCGAGCGGATCACTCAAAAGCAA 1548
858 .....ValAlaThrThrGlyAlaThrLeuThrThrThrAla 869
1549 GTTGATGATCTTATCGCAAAAGTACGCGCAAAATTCACCAAGATGAGCT 1598
870 SerThrAsp.....SerGlyGlySerGly..... 877
1599 ATCAAAAGTTGTGATACATGATGCTCAAAATGCTCAAAATGCA 1648
878 .....LeuAlaGlyTyrGluValLeuArg.....ValS 887
1649 CAACACAGCTTAGATTAATGATATCTCATCTGTAAGTGCATTACC..... 1692
887 erGlyThrThrGlnThrLeuValAlaSerProThrThrAlaThrValAla 903
1693 .....TCGTC 1697
904 LeuAlaGlyLeuThrProAlaThrAlaTyrSerTyrValValArgAlaLys 920
1698 TTAATGATTCGAGAAATGATTA.....GTGCGTCCACACTT 1732
920 sAspGlyAlaGlyAsnValSerAlaValaLysSerProValThrThrThr 937
1733 CAATGTTGATCAAGATTATCTCTCTTCATTAATTTCTAGGGGATCTCAG 1782
937 hrLeuProValThrSerThrProSerGlyThrValValTyrSerThrAsn 953
1783 CATTTGAGCTACGCGCTGCGCGCTGCGAGCGGTTCTCAAGATTGGACCTA 1832
954 SerTyrAsnValGlyPhe.....ThrGlySerValLysIleThr.. 966
1833 CGGCGCTCGGCGCGGCTGCTACCGACGATTCAGCTACGCGCTGCGGCC 1882
967 .....AsnThrGlyThrThrProLeuThrThrThrLeuGlyPheAla 981
1883 CTGCGCAGCGGTAGCCAGATTGAGCTACGCGCTCGCGTCCGCGTGAATCT 1932
981 heProSerGlyGlnGln.....ValThrGlnGlyTyrPser 992
1933 CAGCATTGAGAGCTAC...GGCTGCGCGCTGCGAGCGGTTCTCAAGATTG 1979

```

```

993 AlaThrTyrPserGlnThrGlyThrThrValThrAlaThrGlyLeuSerTyr 1009
1980 GAGCTACGCGCTGCGCGGCTGCTCTACGACGATTCGAGCTACGCGC 2029
1009 pAsnAlaThrLeuGlnProGlyGlnSerThrAspIleGlyPheAsnLys 1026
2030 TGGCGCCCTGCGAGCGGTAGCCAGATTGAGCTAC 2064
1026 erHisProGlyThrAsnThrAsnProAlaSerPhe 1037

seq_name: SwissProt_39:TALI_MOUSE

seq_documentation_block:
ID TALI_MOUSE STANDARD; PRT; 2541 AA.
AC P26039;
DT 01-MAY-1992 (rel. 22, Created)
DT 01-MAY-1992 (rel. 22, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE TALI.
GN TLN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-FIBROBLAST;
RX MEDLINE-91015390; PubMed-2120593;
RA Rees D.J.G., Ades S.A., Singer S.J., Hynes R.O.;
RT *Sequence and domain structure of talin.*;
RL Nature 347:685-688(1990).
CC -!- FUNCTION: PROBABLY INVOLVED IN CONNECTIONS OF MAJOR CYTOSKELETAL
CC STRUCTURES TO THE PLASMA MEMBRANE. HIGH MOLECULAR WEIGHT
CC CYTOSKELETAL PROTEIN CONCENTRATED AT REGIONS OF CELL-SUBSTRATUM
CC CONTACT AND, IN LYMPHOCYTES, AT CELL-CELL CONTACTS.
CC -!- SUBUNIT: BINDS WITH HIGH AFFINITY TO VINCULIN AND WITH LOW
CC AFFINITY TO INTEGRINS.
CC -!- PTM: PHOSPHORYLATED.
CC -!- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,
CC RADIXIN, AND TALIN.
CC -!- SIMILARITY: IN THE C-TERMINUS TO YEAST SLA2 AND C.ELEGANS ZK670.3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X56123; CAA39588.1; -
CC PIR; S11661; S11661.
CC HSSP; P04002; 1WFA.
CC MGD; MGT:1099832; TLN.
CC INTERPRO; IPR000299; -
CC INTERPRO; IPR002358; -
CC PFAM; PF00373; Band_41; 1.
CC PFAM; PF01608; I_LMED; 1.
CC PROSITE; PS00660; BAND_41_1; 1.
CC PROSITE; PS00661; BAND_41_2; 1.
CC PROSITE; PS00657; BAND_41_3; 1.
CC Structural protein; Cytoskeleton; Phosphorylation.
CC FT DOMAIN 173 332
CC FT VARIANT 1105 1105 L -> P.
CC FT VARIANT 2180 2180 K -> M.
CC SO SEQUENCE 2541 AA; 269832 MW; 14EF5ABE9FC2CB CRC64;

alignment_scores:
Quality: 159.00 Length: 902
Ratio: 0.408 Gaps: 39
Percent Similarity: 43.237 Percent Identity: 20.067

alignment_block:

```

US-09-306-689-12 x TALI\_MOUSE ..  
Align seg 1/1 to: TALI\_MOUSE from: 1 to: 2541

```

35 GCACGGCCTGCGCCCTGGCAGCGGTTCTCAGATTGAGACTAGCGCTG 84
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
706 AlAThGlnCysAlaLeuSerThrSerGlnLeuVal.....AlaCy 719
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
85 CGCGCGGGGCTCTAGCCAGCTTGAGCTAGCGGCTGC.....GCC 128
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 sThrLysValAlaAlaProThrLieserSerProValCysGlnGlnL 736
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 TGGCAGCGGTAGCCAGATTGAGCTAGCGGC.....TGCCTCGG 169
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
736 euValGlnLagLysArgLeuValAlaLysAlaValGlnGlyCysVal 752
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
170 GTGAGTCTCAGC.....ATTGGAGC 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
753 AlaserGlnAlaLathrGlnAspGlyGlnLeuLeuArgGlyValGly 769
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
190 TACGGCTCGCCCTGGCAGCGGTTCTC.....A 218
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
769 aAlaAlaThrAlaValThrGlnAlaLeuAsnGlnLeuGlnHisValL 786
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 AGATTGAGCTAGCGGCTGC..... 238
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
786 yAlaHisAlaThrGlyAlaGlyProAlaGlyArgTyrAspGlnAlaThr 802
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
238 ..... 238
803 AspThrIleLeuThrValThrGlnAsnIlePheSerSerMetClyAsp 819
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 .....GTCCGGTGGCTCTAGCCAGCATTTGAGCTAGCGGCTCGG 278
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
819 aGlyGlnMetValArgGlnAlaArgIleLeuAlaGlnAlaThrSerAsp 836
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
279 CCCTGGCAGCGGTA..... 292
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
836 euValAsnAlaIleLysAlaAspAlaGlnGlyGlnSerSplLeuGln 852
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 ...GCCAAGATTGAGCTAGCGGCTCGGTCGCGGAGATTGATTCGCA 339
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
853 SerArgLysLeuLeuSerAlaAlaLysIleLeuAlaAspAlaThrAla 869
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
340 AAAACGTGGGCAAAAAATTTATCCTATATTCCCA.....A 379
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
869 sMetValGlnAlaAlaLysGlyAlaAlaAlaHisProAspSerGlnG 886
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
380 ATTACCAATATGATCTACTGAACAAGTAATGTTTACAGATTATTCACA 429
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
886 lngLngLthrGlyLeuArgGlnAlaAlaGlnGlyLeuArgMetAlaThr 902
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
430 GCGCGCGAAGAGCTGGGAGTTGAG.....GTACAAGAAGAAAGAC 470
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
903 AlAlaAlaLagLnsAlaAlaIleLysLysLysLeuValGlnArgLeu 919
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
471 C.....ATAATATTGCAACAGCTC 490
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
919 sAlaAlaLysGlnAlaAlaAlaSerAlaLathrGlnThrIleAlaAla 936
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
491 AAACAGTTTAGCAGATTCAACCGCTATTGGC.....TTAACT 531
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
936 lnhAlaAlaLaserAlaProLysAlaSerAlaGlyProGlnProLeu 952
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
532 GAGCGTGCATTGTATTCCGCTCCACAATTTGATTAATTGCTACAGA 581
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
953 ValGlnSerCysLysAlaValAlaGlnGlnIleProLeuLeuValGln 969
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
582 AACTAAAGACGCCAAGCA..... 600
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
969 yValArgGlySerGlnAlaGlnProAspSerProSerAlaGlnLeuAl 986
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 ...TTAGGTCTCGCAAAAGCATTTGACAAATGCAAAATAA..... 639

```

```

986 euLeAlaLaserGlnSerPheLeuGlnProGlyLysMetValAla 1002
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
640 ...GCCAAACGTATTATCTGGCATTCAA..... 666
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1003 AlAlaAlaLysAlaSerValProThrIleGlnAspGlnAlaSerAlaMet 1019
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
667 .....TCTATTTTAGCTCAGTATTGGCTGGAATGAGAT 700
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1019 nLeuSerGlnCysAlaLysAsnLeuGlyThrAlaLeuAla.....G 1034
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
701 TAGATGAGCCTTACAGATTAAACAGCAACCAATGCTGCTGCTAAAGCT 750
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1034 euArgThrAlaAlaGln.....LysAlaGlnGlnAlaCysGlyProLeu 1048
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 GGCTTGAGCTAACAAATTCATTAAATTTGAAATTTGCTAATTCAGTAA 800
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1049 GlnMetAspSerAlaLeuSerValValGlnAsnLeuGlyLysAsp.... 1063
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
801 AACACTTGACGAATTTGGTGACCAATTAGTCAATTTGGTCAAAACTAC 850
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1064 .....LeuG 1065
851 AAAATATCAAGCTTAGGACTTTAGAGACAACTCAAAATATTCGCT 900
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1065 lngLnlLeuLysAlaAlaAlaArgAspGly...LysLeuLysProLeu... 1079
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
901 GGACTTGATMAAGCTGGCCTTGTTAGATTATCTCAGGCGCTATTATC 950
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1079 ..... 1079
951 GGGCGCAACAGCTGCATTGTACTTGACAGATATAAATGCTTCAACAGCTA 1000
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1079 ..... 1079
1001 AAAAAGTGGTGCGGGTTTGAATTGGCAACCAAGTTGTGTATATTT 1050
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1080 ...ProGlyGlnThrMetGlnLysCysThrGlnAspLeuLysSer 1094
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1051 ACCAAAGCCGTTTCTTCTTAC..... 1071
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1095 ThrLysAlaValSerSerAlaIleAlaLysLeuGlnGlyLleAlaG 1111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1072 .....ATTTAGCCCAACAGCTGTGACAGAGTT 1099
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1111 nGlyAsnGlnAsnTyrAlaGlyIleAlaAlaArgAspValAlaGlyL 1128
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1100 TATCTTCAACTGGGCTGCGCTTAAATTGCTTCT...ACTGTTCT 1146
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1128 euArgSerLeuAlaGlnAlaAlaArgGlyValAlaAlaLeuThrSer 1144
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1147 CTTCGATTAGCCCATTTAGCATTTGCCGATTATGCCGATTAATTAATCA 1196
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1145 ProAlaValGlnAlaIleValLeuAspThrAlaSerAspValLeuAsp 1161
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1197 TGCAAAAGTTTAGAGAGTTATGCCGACGCTTTAAAAAATTAGGCTATG 1246
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1161 sAlaSerSerLeuIleGlnGlnAlaLys...LysAlaSerGlyHisP 1176
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1247 ACGGATTAATTTATTAGCAGATATTCAGCGGGGAACAGGCGCTATTCAT 1296
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1176 roGlyAsp.....ProGlnSerGlnGlnArgLeuAlaGlnAla 1189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1297 GCATCGGTTACTGCA..... 1311
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1190 LysAlaValThrGlnAlaLeuAsnArgCysValSerCysLeuProGly 1206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1312 .....ATTAAATACCGCATTTGGCGCTATTGCTGTGTGTGCTG 1354
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1206 nArgAspValAspAsnAlaLeuArgAlaValGlyLysPalaSerLysArg 1223
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1355 CTGCAGCGGATTTA.....ACATTTGAAAAGTAA 1386
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

1223 euleuSerAspLeuProSerThrGlyThrPheGlnGluAlaGln 1239
1387 CATAACTTTCATCAGCAATAGCAAAAAAGAAAGTGAACATTCAAA 1436
1240 SerArgLeu..... 1242
1437 CTGGTTCGAGAGCGTGATTGCTTAAGAAAGTCCTAATTAATACAA 1486
1243 .....AsnGluAla.....AlaAlaG 1248
1487 CTAAAGATGCAAAATCAAGAAATCATCGGTCAAAATGCGAGCGATC 1536
1248 LysLeuSngInAlaAlaThrGlnLeuValGln..... 1258
1537 ACCCAAGCAAGAGTGTATGATCTAATGCAAAAGTAAAGCAAAATTAC 1586
1259 AlaserArgLysThrProGlnAspLeuAlaArgAlaSerGlyArgPheGln 1275
1587 CCAAGAT.....GAGC 1597
1275 YGlnAspPheSerThrPheLeuGlnAlaGlyValGlnMetAlaGlyGln 1292
1598 TAACAAGAGTTGTTGATTAAGTAAATGCTCAAAACATAGCAAAATGTG 1647
1292 LaproSerGlnGluAspArgAlaGlnValSerAsnLeuLysGlyLe 1308
1648 ACAAAAGCTTAGATAGTAAATCTCATCTGTAAGTCAATTACCTGCTC 1697
1309 SerMetSerSerLysLeuLeuAlaAlaLysAlaLeuSerThrAs 1325
1698 TAATGATTCAGAAATGTA.....TTAGTGGCTCCAACTTCAATGT 1738
1325 ProAlaSerProAsnLeuLysSerGlnLeuAlaAlaAlaArgAlaVal 1342
1739 TGATCAAAATTTATCTCTCTT.....CAATTGGCTAGG 1773
1342 alThrAspSerLLeAsnGlnLeuLethrMetCysThrGlnAlaPro 1358
1774 GGAATCTAGATTGAGCTACGCGCTCGC..... 1803
1359 GlnGlnLysGlnLucyAspAsnAlaLeuArgGlnLeuLthrValArgGln 1375
1804 .....CTGGACAGCGGTTCTCAAGATTGAGCTAC...GGCC 1837
1375 uLeuLeuGlnAsnProValGlnProIleAsnAspMetSerLysPheGly 1391
1838 TGGCTCCGGGTGGCTACGACG.....ATTGAGCTACGCGCTG 1878
1392 CysLeuAspSerValMetGlnAsnSerLysValLeuGlnAlaMetThr 1408
1879 CGCCCTGGACGGGTAGCCAAAGATTGA..... 1906
1408 rGlyIleSerGlnAsnAlaLysAsnGlnLysLeuProGlnPheGlyAsp 1425
1907 .....GCTACGCGCTCGCTCGGCTGATCTCAGCAATTGAGCTACGGCC 1951
1425 LalaLeaLthrAlaSerLysAlaLeu.....CysGlyPheThrGln 1438
1952 TGGCCCTGGACGGGTCTCAAGATTGA..... 1991
1439 AlaAlaAlaGlnAlaAlaLysLeuValGlyValSerAspProAsnSerGln 1455
1982 .....GCTACGG 1988
1455 nAlaGlyGlnGlnLeuValGlnProThrGlnPheAlaArgAlaAsnG 1472
1989 CTTGGCTCCGGGTGCTTAGCCAGCATTTGAGCTACGCGCTGC..... 2032
1472 LmaLarIleGlnMetAlaCysGlnSerLeuGlnProGlyCysThrGln 1488
2033 .....GCCCTGGACGGGTAGCCAAAGATTGAGCTACGGC 2067
1489 AlaglnValLeuSerAlaAlaLthrIleValAlaLysHisThrSerAlaLe 1505

```

2068 CTGC 2071  
 111  
 1505 uCys 1506

seq\_name: SwissProt\_39:FLIC\_SALON

seq\_documentation\_block:

ID FLIC\_SALON STANDARD; PRT; 507 AA.

AC 006974:

DT 01-OCT-1996 (rel. 34, Created)

DT 01-OCT-1996 (rel. 34, Last sequence update)

DT 30-MAY-2000 (rel. 39, Last annotation update)

DE FLAGELLIN (PHASE-1-C FLAGELLIN).

GN FLIC.

OS Salmonella oranienberg.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Salmonella.

RP (1)

RC SEQUENCE FROM N.A.

RX STRAIN-ATCC 9239; MEDLINE=93374829; PubMed=7690024;

RA Masten B.J., Joys T.M.;

RT "Molecular analyses of the Salmonella g... flagellar antigen

RT complex.";

RL J. Bacteriol. 175:5359-5365(1993).

CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO

CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.

CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE

CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED

CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

DR EMBL: 215070; CAA78779.1; -  
 DR INTERPRO: IPR001029; -  
 DR INTERPRO: IPR001492; -  
 DR PFAM: PF00700; Flagellin\_C; 1.  
 DR PFAM: PF00669; Flagellin\_N; 1.  
 DR PRINTS: PR00207; FLAGELLIN.  
 KW Flagella.  
 FT INIT\_MET 0  
 FT SEQUENCE 507 AA; 53065 MW; 17A88B4BBA4A6125 CRC64;  
 SQ

alignment\_scores:

Quality: 157.50 Length: 501

Ratio: 0.592 Gaps: 23

Percent Similarity: 53.094 Percent Identity: 21.557

alignment\_block:

US-09-306-689-12 x FLIC\_SALON ..

Align seg 1/1 to: FLIC\_SALON from: 1 to: 507

```

469 CGCAATATATTGCAACAGCTTAACAGGTTAGGACGATTCACACCGC 518
14 GlnAsnAsnLeuAsnLysSerGlnSerLeu.....SerSerAl 27
519 TATGGCTTAAGTACGAGCGGTGATGTTATCGCTCACCAATTTGATA 568
27 alleGluArgLeuSerSerGlyLeuArgIleAsnSerAlaLysAsp 43
569 AATTGCTACGAAACAACTAAAGCAGGCAACGATTA..... 603
44 .....AlaAlaGlyGlnAlaAlaIleAlaAsnArgPheThr 54

```

```

604 GGTCTGCCGAAAGCATTTGACAA.....AATGCAATATA..... 639
      : : : : : : : : : : : : : : : : : : : : : : : :
55 SerAsnIleLysGlyLeuThrGlnAlaSerArgAsnAlaAsnAspIleYl 71
      : : : : : : : : : : : : : : : : : : : : : : : :
640 .....GCCAAACTGTATTATCTGGCATTCATCTATTATTTAGGCTCAG 682
      : : : : : : : : : : : : : : : : : : : : : : : :
71 eSerIleAlaGlnThrThrGluGlyAlaLeuAsnGlnIleAsnAsnAsn 88
      : : : : : : : : : : : : : : : : : : : : : : : :
683 TATTGGCTGCATGCAATTAGATGAGCCCTTACAGAAATACAGCAACCA 732
      : : : : : : : : : : : : : : : : : : : : : : : :
88 euGlnArgValArgGluLeuSerValGlnAlaThrAsnGlyThrAsnSer 104
      : : : : : : : : : : : : : : : : : : : : : : : :
733 CATGCTCTGCTAAAGCTGCTGGAGCTACAAATTCATTATTAATTAATA 782
      : : : : : : : : : : : : : : : : : : : : : : : :
105 AspSerAspLeuLysSerIleGlnAspGluIleGlnGlnArgLeuGlu 121
      : : : : : : : : : : : : : : : : : : : : : : : :
783 TATT.....GCTAATTCGATAAAACACTTG 808
      : : : : : : : : : : : : : : : : : : : : : : : :
121 uIleAspArgValSerAsnGlnThrGlnPheAsnGlyValLysValLeu 138
      : : : : : : : : : : : : : : : : : : : : : : : :
809 AGCAATTTGCTGAGCAATTAAGTCAATTGTTCAAAA..... 846
      : : : : : : : : : : : : : : : : : : : : : : : :
138 exGlnAspAsnGlnMetLysIleGlnValGlyAlaAsnAspGlyIuThr 154
      : : : : : : : : : : : : : : : : : : : : : : : :
847 .....CTACAA.....AATATCAAGGCTTAGG..... 870
      : : : : : : : : : : : : : : : : : : : : : : : :
155 IleThrIleAspLeuGlnLysIleAspValLysSerLeuGlyLeuAsp 171
      : : : : : : : : : : : : : : : : : : : : : : : :
871 .....ACTTTAGAGAC.....A 883
      : : : : : : : : : : : : : : : : : : : : : : : :
171 yPheAsnValAsnGlyProLysGluAlaLeuThrValGlyAspLeuLys 188
      : : : : : : : : : : : : : : : : : : : : : : : :
884 AACTCAAAAATATCGTGACATTGATAAGCGCTTGCT..... 924
      : : : : : : : : : : : : : : : : : : : : : : : :
188 exPheLysAsnValThrGlyTyArgPThrTyAlaValGlyAlaAsnLys 204
      : : : : : : : : : : : : : : : : : : : : : : : :
925 .....TTAGATGTTATCTCAGGCGCTTATCGCGCAACAGCTGCACAT 968
      : : : : : : : : : : : : : : : : : : : : : : : :
205 TyArgValAspValAsnSerIleValAlaValThrAspThrThrAla 221
      : : : : : : : : : : : : : : : : : : : : : : : :
969 TGTACTTGCGATATAA.....AATGCTTCAACAGCTATAAAAGTGC 1009
      : : : : : : : : : : : : : : : : : : : : : : : :
221 cThrValProAspLysValTyValAsnAlaAlaAsnGlyGlnLeuThr 238
      : : : : : : : : : : : : : : : : : : : : : : : :
1010 GTGGCGGTTTGAATTGGCAACCAAGTTGTTGTAATTAATCAACAGCC 1059
      : : : : : : : : : : : : : : : : : : : : : : : :
238 hrAla.....AspAlaGlnAsnAsnThrAlaValAspLeuPheLysSer 252
      : : : : : : : : : : : : : : : : : : : : : : : :
1060 GTTCTCTTACATTTTAGCCCAAGCTGTTGACAGAGCTTATCTTCAAC 1109
      : : : : : : : : : : : : : : : : : : : : : : : :
253 ThrLysSer.....AlaAlaGlyThrAspAspAl 262
      : : : : : : : : : : : : : : : : : : : : : : : :
1110 TGGGCGCTGCTGCTTAAATGTTCTTACTGTTCTTCTGCGATTAGCC 1159
      : : : : : : : : : : : : : : : : : : : : : : : :
262 aLysAlaIleAlaThrSerIleLysGly..... 271
      : : : : : : : : : : : : : : : : : : : : : : : :
1160 CATTTAGCATTTGCGGATTCGCCGATAATTAATCAACCAAAAGTTTA 1209
      : : : : : : : : : : : : : : : : : : : : : : : :
272 .....GlyLysValGlyAspThrPheAsp..... 279
      : : : : : : : : : : : : : : : : : : : : : : : :
1210 GAGAGTTATGCGCAAGCTTAAAAAATTAGCTTGAACGAGATTAATT 1259
      : : : : : : : : : : : : : : : : : : : : : : : :
280 .....TyTrpSerGlyValSerPheThrIleAspThr 290
      : : : : : : : : : : : : : : : : : : : : : : : :
1260 ATTACAGAAATATCAGCGGGAACAGGACATTAATGATGATCGTTACTG 1309
      : : : : : : : : : : : : : : : : : : : : : : : :
290 salAcGlyAsp.....AspGlyAsnGlyThrValSerThrThrIleAsnG 305
      : : : : : : : : : : : : : : : : : : : : : : : :
1310 CAATTAAATACCGCATTTGCCGCTTATGCTGCGTGCTGCTCTCTGCA 1359
      : : : : : : : : : : : : : : : : : : : : : : : :
305 yGluLysValThrLeuThr.....IleSerAspIleGlyAlaSerAla 319

```

```

1360 GCCGATTTAACATTTGAAAAAGTTAAACATATCTTGTCAATCAAGATAG 1409
      : : : : : : : : : : : : : : : : : : : : : : : :
320 ThrAspValAsnSerAlaLysIleGln.....Se 329
      : : : : : : : : : : : : : : : : : : : : : : : :
1410 CAAAAAGAGAAAGTGACC.....ATTCAAAACTGTTCCGAGAGCTG 1453
      : : : : : : : : : : : : : : : : : : : : : : : :
329 rSerLysAspValTyThrSerValAlaSerGlyGlnPheThrPhAla 346
      : : : : : : : : : : : : : : : : : : : : : : : :
1454 ATTTGCTTAA.....GAAGTGCCTTATTATAAGCACTATAA 1491
      : : : : : : : : : : : : : : : : : : : : : : : :
346 sPlyThrLysAsnGlnSerAlaLysLeuSerAspLeuGluAlaAsn 362
      : : : : : : : : : : : : : : : : : : : : : : : :
1492 GATGAGAAATGCAAGAAATTCGCTCAAAATGCGGAGCGATCACTTC 1541
      : : : : : : : : : : : : : : : : : : : : : : : :
363 AlAlaValLysGlyGluSerLysIleThrValAsnGlyAlaGluThrAl 379
      : : : : : : : : : : : : : : : : : : : : : : : :
1542 AAAGCAAGTTGATCTTATCGCAAAAGTAAAGCGCAAAATTAACCAAG 1591
      : : : : : : : : : : : : : : : : : : : : : : : :
379 aAsnAlaIaGlyAspLysValThrLeuAlaGlyLysThrMetPheIle 396
      : : : : : : : : : : : : : : : : : : : : : : : :
1592 ATGAGCTATCAAAAGTTGT.....GATAACTATGAA 1623
      : : : : : : : : : : : : : : : : : : : : : : : :
396 sPlyThrAlaSerGlyValSerThrLeuIleAsnGluAspAlaAla 412
      : : : : : : : : : : : : : : : : : : : : : : : :
1624 TTGCTCAACATAGCAAAATGTGACAAACAGCTTATGATTAATCTTC 1673
      : : : : : : : : : : : : : : : : : : : : : : : :
413 AlAlaLysSerThrAlaAsnProLeuAlaSerIleAspSerAlaLeu 429
      : : : : : : : : : : : : : : : : : : : : : : : :
1674 ANCTGTACTGCAATTTACCTCGCTAATGATGTCG...AGAAATGTATTA 1720
      : : : : : : : : : : : : : : : : : : : : : : : :
429 rLysValAspAlaValArgSerSerLeuGlyAlaIleGlnAsnArgPhe 446
      : : : : : : : : : : : : : : : : : : : : : : : :
1721 TGGCTCAACTTCATGTTGATCAAAAGTTTATCTTCTTCAATTGCT 1770
      : : : : : : : : : : : : : : : : : : : : : : : :
446 sPserAlaIleThrAsnLeuGlyAsnThrValThrAsnLeuAsnSerAla 462
      : : : : : : : : : : : : : : : : : : : : : : : :
1771 AGG 1773
463 Arg 463

```

```

seq_name: SwissProt_39:Y109_YEAST
seq_documentation_block:
ID Y109_YEAST STANDARD: PRT: 995 AA.
AC P40442.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOPHETICAL 99.7 KDA PROTEIN IN SDL 5'REGION PRECURSOR.
GN Y11169C OR Y19402.07C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., DeJin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jørgels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.,
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
DR EMBL: Z46921; CAA87023.1; -

```

```

1069 .....TACATTAG 1078
294 aThrAlaSerAsnSerLeuSerSerSerAspGlyThrIleTyrLeuProT 311
1079 CCCAAGCTGTTCAGCAGAGGTTATCTTCACACTGGCCGTGGCTCT... 1125
    ::::::::::::::::::::| | | | | | | | | |
311 hrThrThrIleSerGlyAspLeuThrLeuThrGlyLysValIleAlaThr 327
1126 .....TTAATGCTTCTACGTTCTCTCTGCGATTACCCATTAGC 1166
    ::::::::::::::::::::| | | | | | | | | |
328 GluGlyValValIleAlaIleGlyAlaLysLeuThrLeuLeuAsp.... 342
1167 ATTGGCGGTATTCGCGAATAATTAAATCAGCAAAAAAGTTTAGAGGTT 1216
    ::::::::::::::::::::| | | | | | | | | |
343 .....GlyAspLysTyrSerPheSerAlaAspLeuLysValT 355
1217 ATGCCGAA.....CGCTTAAATAATAGCCTTAGCGAGAT... 1254
    ::::::::::::::::::::| | | | | | | | | |
355 yrcLyspLeuLeuValLysLysSerLysGluThrTyrProGlyThrGlu 371
1255 ...AATTATTTACAGAAATATCAGCGGGGACACAGGACTATTGATGCATC 1301
    ::::::::::::::::::::| | | | | | | | | |
372 PheAspIleSerGlyGluAsnPheAspValThrGlyAsnPheAsnIleGl 388
1302 GGTACTGCAATTAATACGCA...TTGGCGGCTATTCGTGGTGGTGT 1348
    ::::::::::::::::::::| | | | | | | | | |
388 uGluSerAlaIleThrSerAlaSerIleTyrSerPheThrProSerSerP 405
1349 CTGCTCTCAGCCCGATTAACATTTGAAAAAGTTAAACATATACCTTCTC 1398
    ::::::::::::::::::::| | | | | | | | | |
405 heAspAsnSerGlyLysPle.....SerLeuSer 414
1399 ATCAGCAATAGCAAAAACAGAAAGTACC..... 1428
    ::::::::::::::::::::| | | | | | | | | |
415 LeuSerLysSerLysGlyValValThrPheSerProTyrSerAsnSe 431
1429 .....ATTCAAA 1435
431 rGlyAlaPheSerPheSerAsnAlaIleLeuAsnGlyLysSerValSerG 448
1436 ACTGATTCGAGAGCGATGTTTGGCTAAAGAAGTCCCTAATTATTAAGCA 1485
    ::::::::::::::::::::| | | | | | | | | |
448 LysGluIleThrGlyAspAspThrGluGlySerValAsnAsnGlyGluIle 464
1486 ACTAAAGATGAGAA.....ATCGAAGAATATCGGTCA 1520
    ::::::::::::::::::::| | | | | | | | | |
465 AsnLeuAspAsnGlySerThrTyrValIleValGluProValSerGly 481
1521 AATGGCAGCGGATCACTCAACCAAGTT..... 1551
481 sGlyThrValAsnIleIleSerGlyAsnLeuTyrLeuHisTyrProAspT 498
1552 .....GATGATCTTATCGCAAAAAGTAACGCGCAAAATTAACCAA 1590
    ::::::::::::::::::::| | | | | | | | | |
498 hrPheThrGlyGlnThrValValPheLysGlyGluGlyValLeuAlaVal 514
1591 GATGAGCTATCAAAAAGTTGTTGATTAACATGAAATTCCTCAAAATACCA 1640
    ::::::::::::::::::::| | | | | | | | | |
515 AspProThrGlu..... 518
1641 AATATGACAACACGCTTAGATTAATTTATCTCATCTTAGTGCAATTTA 1690
    ::::::::::::::::::::| | | | | | | | | |
519 .....ThrAsnAlaThrPro.....IleProValValLysTyrT 530
1691 CCGCGCTATATGTTCCGAGAATGTATTAGTGGCTCAACTTCAAATGTTG 1740
    ::::::::::::::::::::| | | | | | | | | |
530 hrClyLysAsnGlnIleAlaIleThrAlaAspIleThrAlaLeuSerTyr 546
1741 GATCAACATTTATCTTCTTCATTAATTTGCTAGGGGATCTCAGATTGGAG 1790
    ::::::::::::::::::::| | | | | | | | | |
547 AspGlyThrThrGlyValLeuThrAlaThrGlnGlyAsnArgInPheSe 563
1791 CTAGCGCTCGGCCCTCGGACGGGTTCTCAAGATTGGAGCTAC.....G 1834

```













## alignment\_block:

US-09-306-689-12 x FLA3\_CAMJE ..

Align seg 1/1 to: FLA3\_CAMJE from: 1 to: 574

```

137 GTAGCCAGATTGGAGCTACGCGCTCGTCCGGGTGATCTC..... 178
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
      8 ValAlaAlaLeuAsnAlaLysAla..... AsnAlaAspLeuAsnSerIly 22
179 .AGCATTTGGAGCTACGCGCTCGCGCGGTTCACAGATTGGAG 227
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
22 sSerLeuAspAlaSerLeuSerArgLeuSerSerGlyLeuArgIleAsnS 39
228 CTACGGCTCGCTCGCGGTGCTCTAGCC..... AGCATTTGGAGCT 268
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
39 erAlaAlaAspAlaSerAlaSerGlyMetAlaIleLysAspSerLeuArgSer 55
269 ACGGCTCGCGCTCGCGAGCGGTAGCCAGATTGGAGCTACGCGCTGCGT 318
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
56 GluAlaAsnThrLeuGlyGlnAlaIleSerAsnGlyAsnAspAlaLeuG 72
319 CCGGGTGGATCTAGCTTCCCAAAACTGGGGCAA..... AAAAATTAT 362
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
72 yIleLeuGlnThrAlaAspLysAlaMetAspGluGlnLeuLysIleLeuA 89
363 CCTCTATTATCCCAAAATTACCAATTGATGACTGACACAGTAACTGTT 412
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
89 spThrIleLysThrLysAlaThrGlnAla..AlaGlnAspGlyGlnSerI 105
413 TACAGGATTTAGTCAAAGCGCGGAGAGTTGGGATTTAGAGTACAAAGA 462
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
105 eu..... LysThrArgThrMetLeuGlnAlaAspIleAsnArg 117
463 ..... GAAGACGCAATATATATGCA..... 483
118 LeuMetGluGlnLeuAspAsnIleAlaAsnThrThrSerPheAsnGly 134
483 ..... 483
134 sGlnLeuLeuSerGlyAsnPheIleAsnGlnGluPheGlnIleGlyAla 151
484 ..... ACAGCTCAAAACGATTAGCAGCAGATTCAACCGCT... 519
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
151 erSerAsnGlnThrValLysAlaSerIleGlyAlaThrGlnSerSerLys 167
520 ATGGCGCTTAAC..... GAGCGTGCATGTGTATCCGCTCA..... 558
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
168 IleGlyLeuThrArgPheGluThrGlySerArgIleSerValGlyGly 184
559 ..... CAAATTG 565
184 uValGlnPheThrLeuLysAsnTyraSnglyIleAspAspPheLysPheG 201
566 ATAAATGCTACAGAAAACCTAAAGCAGCCAGCATTAGCTTCT...GCC 612
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
201 InLysValValIleSerThrSerValGlyThrGlyLeuGlyAlaLeuAla 217
613 GAAGAGCTTGTACAAATGCAAAATAAGCCAAACTGATTATCTGTCAT 662
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
218 AspGluIleAsnLysAsnAlaAspLys..... ThrGlyLys 229
663 TCATATCTATT..... TTAGCTCAGTATTGGCTG 691
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
229 IArgAlaThrPheThrValGluThrArgGlyMetGlyAlaValArgLys 246
692 GAATGATTTAGATGAGCGCTTACAGATAACAGCAACCAACTGCTCTT 741
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
246 GylAlaThrSerAspAspPheAlaIleAsn..... 255
742 GCTAAAGCTGCTTGGAGCTA..... AC 764
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
256 ..... GlyValLysIleGlyLysValAspTyrLysAspGlyAspAl 269

```

```

765 AAATTCATTATTAATAATATTCATTACGATAAA..... 801
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
269 aAsnGlyAlaLeuValSerAlaIleAsnSerValLysAspThrThrGly 286
802 ..... ACATTTGACGAAATTGGTGAGCAAT..... AGTCATTT 837
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
286 aGlnAlaSerIleAspGluAsnGlyLysLeuLeuThrSerArgLys 302
838 GGTCAAAACTACAAAATATCAAAAGCTTAGGACTTAA..... 879
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
303 Gly..... ArgGlyIleLysIleGlyLysAsnIleGlyArgGlyAl 316
880 ..... GACAACTCAAAAATATCGGTGACTGATTA... 912
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
316 aPheIleAsnProAsnMetLeuGlnAsnTyrGlyArgLeuSerVal 333
913 ..... GCTGGCTTGGTTAGATGATTATC 936
333 yAsnAspGlyLysAspIleLeuIleSerGlyThrAsnLeuSerAlaIle 349
937 ..... TCAGGCTATTATCGGCGCAACAGCTCAGCTGTACT 974
350 GlyPheGlyThrGlyAsnMetIleSerGlnAlaSerValSerLeuArg 366
975 TGCA..... GATAAAAATGCTTCAACAGCTAAAAAAGTGGTG 1012
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
366 uSerLysGlyGlnIleAspAlaAsnValAlaAspAla..... M 379
1013 CGGGTTTGAATTGGCAACCAAGTTGTTGTAATATACCAAAAGCGCT 1062
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
379 etGlyPheAsnSerAlaAsnLys..... GlyAsnIleLeuGlyGlyTyr 393
1063 TCTTCTTACATTTTAGCCCAAGCTGTGACAGCAGGATTATCTTCACTG 1112
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
394 SerSer..... ValSerAlaTyrMetSerSerThrG 404
1113 G..... 1113
404 ySerGlyPheSerSerGlySerGlyPheSerValGlySerGlyLysAsn 421
1114 ..... CTTGCTGCTGCTTAAATGCTTACTAGT 1143
421 ySerThrGlyPheAlaAsnThrIleAlaIleSerAlaAlaSerGlnLeu 437
1144 TCTCTGCGATTAGCCCATTAAGATTGCGCGTATGCCGATAATTAA 1193
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
438 SerAlaValTyrAsnValSerAlaGlySerGly..... Phe 450
1194 TCATGCAAAAAGTTTAGAGATTATGCCGAACGCTTAAAAATTAGCT 1243
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
450 rSerGlySerAsnLeuSerGlnPheAlaThrMetLysThrSerAlaGly 466
1244 ATGACGGAGATATTTATTAGCA..... GAATTCACGCGGGGACACGG 1287
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
467 ..... AsnThrLeuGlyValLysAspGluThrAlaGlyValThr 479
1288 ACTATTGATGCATCGTTACTGCAATTAAATACCGCATGGCGCTATTGC 1337
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
480 ThrLeuLysGlyAlaMetAlaValMetAspIleAlaGluThrAlaIle 496
1338 TGTGTGTGTGCTGCTGCTGCGAGCTTAACTTTGAAAAGCTTAAC 1367
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
496 r..... AsnLeuAspGlnIleArgAlaAspIleGlySerValGlnAsnGln 512
1388 ATATCTTGTTCATCAGAAATAGCAAAAAGAGAGACAT..... 1431
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
512 aThrSerThrIleAsnAsnIleThrValThrGlnValAsnValLysAla 528
1432 ..... CAAAAGCTGCTCGAGAGCTGATTGCTTAAAGAGCTTAATTA 1478
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
529 AlaGluSerGlnIleArgAspValAspPheAlaAlaGluSerAlaAsnTy 545
1479 TAAAGCAACTAAAGATGAGAAAATCGAAGAATCATCGGTCAAAATGCG 1527

```

545 rserLysala.....AsnrlleuAlAGlnSerCly 555

**THIS PAGE BLANK (USPTO)**

OM of: US-09-306-689-12 to: PIR\_66:\* out\_format : pfs  
Date: Mar 2, 2001 10:36 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODE=frim-np.model -DEV=rlp  
-Q/cgn1\_1/USPTO.spool/US09306689.r/unat\_02032001\_102825\_9560/app\_query.fasta\_1.2389  
-DB=PIR\_66 -PEM=fastan -SPEFT=PIR -GAP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -OGAP=4.500  
-OGAPEXT=0.050 -XGAP=10.000 -XGAPEXT=0.500 -FGAP=6.000  
-FGAPEXT=7.000 -YGAP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blomsun62 -TRANS=human40.cdi  
-LIST=45 -DOCALIN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=2000000000 -USER=US09306689 -CGN1\_1\_370 -NCPU=6 -ICPU=3  
-LONGLOG -NO\_XLPHY -WAIT -THREADS=1

## Search information block:

Query: US-09-306-689-12  
Query length: 2088  
Database: PIR\_66:\*  
Database sequences: 195891  
Database length: 6790655  
Search time (sec): 196.850000

## score list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
PIR1:OYBRC	+	212.00	199.47	0.0002	1706	Documentation
PIR2:S51672	+	206.50	194.17	0.0004	1705	cyclolysin - Bordetella pertussis
PIR2:A57036	+	182.50	168.14	0.0080	2491	adenylate cyclase hemolysin -
PIR1:KRNU2	+	172.00	168.40	0.0301	643	keratin 1, type II, cytoskeletal
PIR2:G93339	+	169.50	149.37	0.0394	5627	hypothetical protein P2462 [A
PIR2:A75564	+	169.00	159.19	0.0430	1467	conserved hypothetical protein
PIR2:A62615	+	168.50	160.31	0.0459	1190	surface protein Xf1981 (import
PIR2:T14961	+	166.50	156.48	0.0585	1525	hypothetical protein Y1044 - Y
PIR2:153169	+	166.00	162.67	0.0632	638	cytochrome 2 - human
PIR2:A44861	+	166.00	162.59	0.0632	645	keratin, 67K type II epidermal
PIR2:E75590	+	166.00	161.37	0.0630	756	methyl-accepting chemotaxis-rel
PIR2:A56143	+	165.00	157.48	0.0708	1109	surface-array protein homolog
PIR2:141061	+	164.00	161.68	0.0811	565	flagellin - Escherichia coli (S
PIR2:137984	+	163.00	159.98	0.0916	622	keratin 9, type I, cytoskeletal
PIR2:T13634	+	163.00	153.19	0.0901	1510	probable minor tail protein gr
PIR2:A53199	+	161.00	154.56	0.1092	1045	endoglucanase B (EC 3.2.1.-)
PIR2:T38274	+	161.00	154.49	0.1163	990	probable transcription initiat
PIR2:A57284	+	160.50	154.80	0.1239	893	surface-array protein - Campylo
PIR2:D82671	+	160.00	147.92	0.1299	2059	surface protein Xf1529 (import
PIR2:T13329	+	159.00	149.30	0.1478	1517	hypothetical protein P4625 [A
PIR2:S11661	+	159.00	145.35	0.1464	2541	tailin - mouse
PIR2:140711	+	158.50	152.51	0.1585	936	sapB protein - Campylobacter fe
PIR2:T34852	+	158.00	150.41	0.1680	1156	probable secreted protein - St
PIR2:S33192	+	157.50	156.24	0.1813	507	phase-1 flagellin - Salmonella
PIR2:A53465	+	157.50	156.23	0.1813	508	phase-1 flagellin - Salmonella
PIR2:R83068	+	157.50	145.17	0.1767	2154	hypothetical protein P4625 [A
PIR2:S50358	+	155.50	149.15	0.2294	995	hypothetical protein Y1169c -
PIR2:R81965	+	155.50	148.61	0.2291	1068	probable outer membrane peptid
PIR2:R81020	+	155.50	148.51	0.2291	1082	serotype-1-specific antigen, F
PIR2:JC1419	+	154.50	152.23	0.2620	587	FC gamma (19c) receptor II pred
PIR2:A47283	+	154.50	149.19	0.2602	873	calphostin - fruit fly (Drosoph
PIR2:T03099	+	154.00	127.86	0.2637	13288	mucin, submaxillary - pig

## seq\_name: PIR1:OYBRC

## seq\_documentation\_block:

cyclolysin - Bordetella pertussis  
N:Alternate names: adenylate cyclase precursor; calmodulin-sensitive adenylate cyclase-  
N:Content: adenylate cyclase (EC 4.6.1.1), calmodulin-sensitive; hemolysin  
C:Species: Bordetella pertussis  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jul-1999

C:Accession: S00893; S14100; S02389

R:Glaser, P.; Ladant, D.; Seeger, O.; Pichot, F.; Ullmann, A.; Danchin, A.

Mol. Microbiol. 2, 19-30, 1988

A:Title: The calmodulin-sensitive adenylate cyclase of Bordetella pertussis: cloning

A:Reference number: S00893; MUID:88216178

A:Accession: S00893

A:Molecule type: DNA

A:Residues: 1-1706 <GLA>

A:Cross-references: EMBL:Y00545; NID:9396665; PIDN:CA68613.1; PID:9396666

R:Glaser, P.; Sakamoto, H.; Bellalou, J.; Ullmann, A.; Danchin, A.

EMBO J. 7, 3997-4004, 1988

A:Title: Sequence of cyclolysin, the calmodulin-sensitive adenylate cyclase--haemoly

A:Reference number: S02386; MUID:89091151

A:Contents: annotation; identification of adenylate cyclase--hemolysin difunctional p

R:Mueller, H.; Gilles, A.M.; Glaser, P.; Krin, E.; Danchin, A.; Sarfati, R.; Barzu, O.

Eur. J. Biochem. 196, 469-474, 1991

A:Title: Isolation and characterization of catalytic and calmodulin-binding domains o

A:Reference number: S14100; MUID:91177021

A:Accession: S14100

A:Molecule type: protein

A:Residues: 1-78, 'W', 80, 'W', 82-97, 'W', 99-139, 'W', 141-178, 'W', 180-399 <MUN>

R:Hackett, M.; Guo, L.; Shabanowitz, J.; Hunt, D.F.; Hewlett, E.L.

Science 266, 433-435, 1994

A:Title: Internal lysine palmitoylation in adenylate cyclase toxin from Bordetella pe

A:Reference number: A55167; MUID:95025937

A:Contents: annotation; lysine palmitoylation

C:Comment: B. pertussis, the etiological agent of whooping cough, disrupts mammalian

action by host cell calmodulin of the adenylate cyclase activity of bacterial cycloly

C:Genetics: Adenylate cyclase activity is activated upon binding of calmodulin in the

A:gene: cyaA, cya

C:Superfamily: cyclolysin; calmodulin-sensitive adenylate cyclase catalytic domain ho

C:Keywords: calcium binding; calmodulin binding; CAMP biosynthesis; carbon-oxygen lya

F:15-326/Domain: calmodulin-sensitive adenylate cyclase catalytic domain homology <AD

F:25-35/Region: calmodulin binding #status predicted

F:54-66/Region: nucleotide binding #status predicted

F:544-1085/Domain: hemolysin A homology <HLY>

F:1033-1041,1042-1050,1174-1182,1289-1297,1298-1308,1316-1324,1430-1438,1556-1564/Reg

F:983/Binding site: palmitate (lys) (covalent) #status experimental

alignment\_scores:

Quality: 212.00 Length: 512

Ratio: 0.763 Gaps: 21

Percent Similarity: 54.297 Percent Identity: 22.266

alignment\_block:

US-09-306-689-12 x OYBRC ..

Align seg 1/1 to: OYBRC from: 1 to: 1706

265 AGCTACGCGCTGCGCCGTCGACGCGGTACCAAGATTGGAGTCGCGCT 314

|||||

349 AATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364

|||||

315 GCGTCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364

|||||

364 AATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381

|||||

365 TCGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414

|||||

381 TATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389

|||||

415 CAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464

|||||

390 ATG.....ATGProserLeuGly...AlaValGluArgG 400

465 AGAAGC.....ATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499

|||||

400 nspSerGlyTyrAspSerLeuAspGlyValGlySerArgSerPheSerL 417

|||||

500 TGGGACGATTCGAACCGCTATGCTTACGAGCGTGGCATTTGCTTA 549

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

[illegible][illegible]

```

seq_name: p1r2:S51672
seq_documentation_block:
adenylate cyclase hemolysin - Bordetella bronchiseptica
C:Species: Bordetella bronchiseptica
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 16-Jul-1999
C:Accession: S51672
R:Betsou, F.; Sismelro, O.; Danchin, A.; Guiso, N.
Submitted to the EMBL Data Library, September 1994
A:Description: The adenylate cyclase-hemolysin gene from Bordetella bronchiseptica.
A:Reference number: S51672
A:Accession: S51672
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1705 <RET>
A:Cross-references: EMBL:Z37112
C:Superfamily: cycloIolysin; calmodulin-sensitive adenylate cyclase catalytic domain homology <AD>
C:Keywords: tandem repeat
F:15-328/Domain: calmodulin-sensitive adenylate cyclase catalytic domain homology <AD>
F:543-1084/Domain:hemolysin A homology <HLTA>

```

```

alignment_scores:      alignment_block:
  alignment_scores:      US-09-306-689-12 x S51672      ..
    Quality: 206.50
    Ratio: 1.059
    Percent Similarity: 52.419
    Length: 3722
    Gaps: 11
    Percent Identity: 23.387

Align seg 1/1 to: S51672 from: 1 to: 1705

```

```

300 TTGGAGCATACGCGCCGCGCTCCGGGAGTACAGTCTCCAAAAACCTGGG 349
      |||||      |||||      ||      |||:      |||
380 IeuGIuThrValProIaSer.ProGIyLeuArgIProSerIeuIya 366
      |||||      |||||      |||||      |||||      |||||
390 CAAAAAAATATCTCTATATCTCCCAAAATTCACATATGATACGAA 399
      ||:      :      |||||      |||||      |||||
396 IavaIglu.....ArgGlnaspSerGIyTyIraspSerIeu 407
      :      :      :      :      :      :      :
400 CAAGGTAATGCTTACAGGATTTATGTCAAACGCCCAAGATGGTGGAT 449
      :|||      |||      :      :      :      :
408 aspGIyValGIySerIasrSerPhe..... 415

```



450 TGAGGTACAAAGAGAAAGCAAGCAATTAATTGGCAACACCTCAACCAAGTT 499  
416 ..... SerL 417  
500 TAGGCACGATTCACAAACCGCTATGTGCTTAACGTAGACGCTGCATTGTGTTA 549  
417 eucllygluValSerSpMet.....AlaAlaVal 426  
550 TCCGCTCCACAAATTGATTAATTGCTACAGAAACTAAAGCAGGCCAGC 599  
427 GluAlaAlaGluLeuGluUmetThrArgGluValLeuHisAlaGlyAlaArg 443  
600 ATTAGTCTTGCCGCAAGACGATTGTACAAAATGCCAATATAA.....GCCA 643  
443 gcGlnspsrAlaGluIleProGlyValSerGlyAlaSerLanHisTrpGlyG 460  
644 AAACGTATTAATCTGCGCATTCATCATTTTATAGCGCTACGTAATGGCTGGA 693  
460 InArgAlaLeuGlnGlnGlyAlaGlnAlaValAlaAlaAla..... 472  
694 ATGGATTTAGATGAGGCGCTTACAGAAATACAGAACCAACAACATCTCTTGC 743  
473 .....GlnArgLeuValHisAlaIleAl 480  
744 T.....AAAGCTGCGCTGGAGCTAAACAATTCATTAA 775  
480 AleUmetThrGlnPheGlyArgAlaGlySerThrAsnThrProGlnIua 497  
776 TTGAAATATTTGCTAATTCAGTAACAAACACCTTGACGAATTTGGTGAGCA 825  
497 LaLaserLeuSerAlaAlaValPheGlyLeuGlyGluAlaLaserSerAla 513  
826 ATTAGTCGA.....TTTGGTCAAAACCTCAAAATATATCAA 860  
514 ValAlaGluThrValSerGlyPhePheArgGlySerSerArgTrpAlaG 530  
861 AGGCTTAGGAGCATTTAGCAGACAAACTCAAAAATATTCGCTGACTGTATA 910  
530 yglPheGlyValAlaGlyGlyAlaMetAlaLeuGlyLeuGlyIleGlyA 547  
911 AAGCTGCGCTGTGTAGATGTATTATCTCAGGCGTATTATCGGCGGCACAA 960  
547 laValGlyAlaGlyMetSer..... 553  
961 GCTGCATTTGATCTTGACAAATGAAATGCTTCAACAGCTAAACAAAGTGG 1010  
554 .....LeuThrAsp..AspAlaProAlaGlyGlnIysAlaAl 565  
1011 TCGGGTTTGAATTGGCAACCAAGTTGTGTATATATTCACAAAGCGC 1060G  
565 aaAlaGlyAlaGluIleAlaLeuGlnLeuThrGlyIleThrValGluLeuA 582  
1061 TTTCCTTCTTCACTTTTA.....GCCACAGTGTTGCAGCAGCT 1098H  
582 lAserSerLleAlaLeuAlaLeuAlaAlaAlaValGlyValThrSerGly 598  
1099 TTATCTTCAACTGAGGCGCTGTGCTGCTTAAATGCTTATCTTACTGTTTCT 1148H  
599 LeuGlnValAlaGlyAlaSerLerAlaGlyAlaAlaAlaGlyAlaLeuAlaAl 615  
1149 TCGGATTAGCCCATTTAGCATTTGCCGGTATTCGCGATTAATTAATCATCG 1198H  
615 aaAlaLeuSerPrometGluIleTyrGlyLeuValGlnGlnSerHisTyrA 632  
1199 CAAAAAGTTTACAGAGTTATGCGCAAGCGTTAAAAAATTTAGGCTATGCAC 1248H  
632 laAspGlnLeuAspLysLeuAlaGlnGlnSerSerAlaTyrGlyTyrGlu 648  
1249 GCAGATAAATTTTATAGCAGATAATCAGCGGGAACAGGACGACTATTTGAGC 1298H  
649 GlysAspAlaLeuLeuAlaGlnLeuTyrArgAspLysThrAlaAlaGlnGln 665  
1299 ATTCGTTACTGCAATTAATATACCGCATGTGGCGCTAATTCGCTGGTGCTGT 1348H

```

:::|||||.....:|||||.....:|||||
665 yalaValAlaGlySerAlaValLeuSerThrValGlyAlaIaValS 682
1349 CTGGCTGCAGCC 1362
|| |||||
682 erlElaIaIaIa 686
seq_name: p1r2:A57036

seq_documentation_block:
talin - slime mold (Dictyostelium discoideum)
C.Species: Dictyostelium discoideum
C.Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 29-Oct-1999
C.Accession: A57036
R.R Reitmeier, M.; Gerisch, G.; Heizer, C.; Mueller-Taubenberger, A.
J. Cell Biol. 129, 179-188, 1995
A.Title: A talin homologue of Dictyostelium rapidly assembles at the leading edge of
A.Reference number: A57036; MUID:95213284
A.Accession: A57036
A.Status: Preliminary
A.Molecule type: DNA
A.Residues: 1-2491 <KRE>
A.Cross-references: GB:U14576; NID:g797298; PIDN:AAC6586.1; PID:g797299
A.Genetics:
A.Introns: 131/1; 222/2

alignment_scores:
quality: 182.50 length: 650
Ratio: 0.583 gaps: 31
Percent Similarity: 48.154 Percent Identity: 22.000

alignment_block:
US-09-306-689-12 x A57036 ..

Align seg 1/1 to: A57036 from: 1 to: 2491

412 TTACAGATTTCACAAGCGGCCGAAGACTGGGATGTAGGTACAAG 461
|||| |::::::::::||| :::::|||
1064 LeuGlyAspLeuValLysSerAlaThrGluSerAlaSerValAlaI 1080
462 AGAAGACGCAATTAATTCGAACAGCTCAACCAGTTTGGCAGATTC 511
::: ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1080 eSerAlaLysThrLeuIleCysAla....ThrThGlyLysGlnValG 1095
512 AAACCGCTAT...GGCTAACGTAGCGGTGCATGTGTTATCCGCTCA 558
|||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1095 InThrLysLeuMetGlyIleThrLysGlnLeuMetIle..... 1107
559 CAATTCATAAATTCGTACAGAAACATAAGCAGCCCAACATTAGGTT 608
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1108 AspMetGlnGlnLeuIleArg.....Al 1115
609 TGCCGAAGCATTTGTACAAATGCAATTAAGCCAACACTGTATTA... 654
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1115 AserArgSerValArgSerAsnProAsnAspArgArgSerGlnLeu 1132
655 .....TCTGCATTCATCTATTTTAGCGTCAGATTTGGCGGA 693
|||| |::::::::::|||
1132 euAspArgSerValArgSerAsnAspValSerIleSerThrAlaAlaLeuValGly 1148
694 .....ATGATTTTGATGAGCGCTTACAGAATAA 722
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1149 SerThrAlaAsnValAspCysLysGlnLeuAspGlnAlaSerAlaSp 1165
723 CAGCAACCAACATGCTCTTGTAAAGCTGGCTTGAGACTTAACAATTC 771
|||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1165 eSerAsnLeuLeuSerLeuLysMetGlySerLeuGlnIleLeuSerG 1182
772 .....TTAATGAAAATNTGCTAATTCAGTA 798
|||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1182 InProThrGlnGluPheAlaPheTyValGlnGlnIleAlaSerThr 1198

```

```

799 AAACACTGACGAATTTGGTACCAAAATTAATGCAATTTGGTTCAAAAC 848
1199 LysAlaLeuAsnAlaAlaSerGlnGlnValValAlaMet...AlaArgAs 1214
849 ACAAAATATCAAGAGCTTAGGACTTTAGGAGACAAACTCAAAATATCG 898
1214 nLysAsnLeuLysGlyLeuGlyAla..... 1222
899 GTGCACTGATAAGCTGGCCTTGGTTAGATTTATCTCAGGCTATTA 948
1223 .....SerAlaLysIleThrAlaSerAlaLeuSerThrLeuVal 1235
949 TCGGGCCCAACAGCTGCATTTACTT.....GCAGATAAAA 986
1236 SerHisAlaGlnAsnAlaIleValLeuThrGlnAsnGlnAlaThrLysAs 1252
987 TGCCTTCAACACTTAAAAAGTGGTGGCGTTTGAATTTGCCAAACCAAG 1036
1252 nAlaIleLeuAlaSerThrValAlaLeuGly.....GlyGlnI 1265
1037 TTGTTGGT..... 1044
1265 LeileGlyLeuLeuAspPheSerLysAlaArgIleAlaAsnTyrLysAsp 1281
1045 .....AATATTACCAAGACCGTTTCTCTTA 1070
1282 ProIleTyrAspGlnAsnLeuIleAsnGlnAlaLysSerValGluAspH 1298
1071 CATTTTACCCCAACGCTGTGCAGCAGTTTATCTTCAACTGGG..... 1113
1298 sLeuVal.....LysValGlyArgSerLeuGlyLysAspLysAsnAsn 1313
1114 .....CCTGGCCTGCTTAATGCTTACTGCTTCTCTT 1149
1313 hrlIeCysAspGluAlaValAlaAspArgIleIleGluAlaThrArgSerLeu 1329
1150 .....GCAGATTAGCCCATTAAGATTGCCGATTCGCCGATAAATTAA 1193
1330 AspLysThrIleLeuPro...AspThrSerGlyLeuGlnThrAsnAlaH 1345
1194 T.....CATGCMAAAAGTTTAAAGACGTTATGCCCAACGCTTAA 1231
1345 sLeuGlnMetLeuHisGlnGlnSerLeuLeuAlaIleThrGlnAlaSerL 1362
1232 AAAAATTAGCTATGACGAGATTAATTATTAAGACAATATCAGCG... 1278
1362 yslLysLeuGlySerIleThrSerAsnLeuValAsnSerLysAsnAsnSer 1378
1279 .....GGACAGGAGACTATTGATCA..... 1299
1379 AspleuValGlySerGlySerThrAspAlaGluArgIleIleGlnMetI 1395
1300 .....TCGCTTACTGCAATTAATACCG 1321
1395 eGluAlaAlaLysHisValAlaHisCysSerIleSerThrTyrAsnProA 1412
1322 CATGGCCCGCTATGCTGTGCTGTGCTGTGCTGCAGCGCATTTAA 1371
1412 spIleLeuLeuProAlaLysSerIleLeuAspAlaSerGlnMetLeuThr 1428
1372 TTTGAAAAAGTTAAACATATCTTGCATCAGCAATAGCAAAAAAGAA 1421
1429 AlaAsnGlnAlaAspValAlaAsnHisValLeuSerHisAla..... 1441
1422 AGTACCAATTCMAAACTGGTCCAGAGAGCTGATTTGCTTAAAGAGTGC 1471
1441 ..... 1441
1472 CTAATTATTAAGCACTAAGATGAGAAATCGAAGAAATCATCGGTCAA 1521
1443 .....AlaThrIleAlaAlaLacYsthrGlnGlnLeuLeuGlyIle 1454
1522 AATGGCAGGCGGATCAC.....TCAAAAGCAAGTTGATGATCTTATCGC 1565

```

```

1455 ThrArgGlnArgAlaSerGlnPheAsnGlnGlnAspGlnGlnGlnValG 1471
1566 AAAAGTAACCGCAAAATTAACCCAGATGAGCTATCAAAAGTTGTGATA 1615
1471 nValArgAspGly.....IleValLys 1479
1616 ACTATGAAATTCCTCAACATATGACAAAAATGACAAACAGCTGATTAAG 1665
1479 erThrGlnGlnLeuAlaHisAla.....ThrSerSerLeuAlaArg 1492
1666 TTAATTCATCTGTAAGTGCATTTACTGCTGTAATGATTCAGACAATGT 1715
1493 AlavalLysSerVal.....ThrSerLysGlnProGlyAlaLysAl 1506
1716 ATTAGTGCTCCCACTTCATGTTGATGATCAAAAGTTTATCTTCTTCAT 1765
1506 aMetIleSer...GlnSerLeuLysAspLeuGlnSerAla...IleAsnAsn 1521
1766 TTGCTAGGAGATCTCAGCATTTGAGCTACG.....GCCTGCGCCTAGC 1809
1522 LeuLeuIleThrSerSerValProAlaSerGlnArgGlyIleGlyIleAl 1538
1810 AGCGGTTCTCAAGATTTGAGCTACGCGCTGCGCGGCTGAGCTTACCA 1859
1538 aAspPheAsnLysLeuMetSerThr...CysArgSerValSerThrAlas 1554
1860 GC.....ATTGACCTACGCGCTGC..... 1879
1554 erSerGlnLeuIleIleSerAlaSerLysSerGlnLysProLysAsp 1570
1880 .....GC 1881
1571 IleGlnLeuSerSerIleLeuSerGlnAsnAlaValLeuMetThrAsnSe 1587
1882 CTGCGACACGCGTACCAAGATTGAGCTACGCGCTGCGCTCCGGTGCATC 1931
1587 rLeuLysAspIleIleLysValThrSerSerMetMetProGlyValAsn 1604
1932 TCACATTTGAGCTACGCGCTGCGCGCTGCGCGCTGCTTCAACATGGA 1981
1604 heCysGlnGluAlaIleIleGlnIleAlaGlnArgAlaIleSerAspLeuSer 1620
1982 GCTACGCGCTGCTGCTGCGGCTGCTTACGCCAGATTTGAGCTACGCGC 2029
1621 SerValAla...LeuSerValAlaValGlySerPheAspSerSerAla 1635
seq_name: p1r1:KRH02
seq_documentation_block:
Keratin I, type II, cytoskeletal - human
N:Alternate names: 67K type II epidermal keratin; cytoKeratin 1
C:Species: Homo sapiens (man)
C>Date: 04-Dec-1986 #sequence-revision 22-Oct-1999 #extLchange 10-Dec-1999
C:Accession: A22940; A02950; A43342
R:Johnson, L.D.; Idler, W.W.; Zhou, X.M.; Roop, D.R.; Steinert, P.M.
Proc. Natl. Acad. Sci. U.S.A. 82, 1896-1900, 1985
A:Reference number: A22940; MUID:85166239
A:Accession: A22940
A:Molecule type: DNA
A:Residues: 1-643 <OH>
A:Cross-references: GB:M98776; GB:M11215; GB:M11845; GB:M11846; NID:g1843461
A:Note: Translation of Initiator Met is not shown
R:Steinert, P.M.; Parry, D.A.D.; Idler, W.W.; Johnson, L.D.; Steven, A.C.; Roop, D.R.
J. Biol. Chem. 260, 7142-7149, 1985
A:Title: Amino acid sequences of mouse and human epidermal type II keratins of M-r 67
late filament subunits.
A:Reference number: A92535; MUID:85207740
A:Accession: A02950
A:Molecule type: mRNA
A:Residues: 151-183, 'K', 185-199, 'M', 201-204, 'K', 206-236, 'S', 238-239, 'R', 241-356, 'Y', 3
'S', 638-643 <STE>
A:Cross-references: GB:M10938; NID:g186787; PIDN:AAA36153.1; PID:g386854

```



```

1177 ATGGCGATAAATTAAATCATGCAAAAAGTTTAGAGAGTTATGCCGACG 1226
      :::::|||||
367 .....GluAlaGluSerLeu.....TyrGlnSerLeu 375
      :::::|||||
1227 CTTTAAAGAAATTA.....GGCTATGACGAGATTAATTATTAG 1264
      :::::|||||
375 sTYrGlnGlnLeuGlnIleThrAlaGlnArgHisGlyAspSerVal.... 390
      :::::|||||
1265 CAGAAATATCAGCGGGGCAACGAGGCTATTGATGCATCGGCTACTGCAATT 1314
      :::::|||||
391 .....ArgAsnSerLeysIleGlnIleSerGlnLeu 400
      :::::|||||
1315 AATACCCGATTGGCGCGCTATTGCTGTGGTGTCTGCTGCTGACGCCGA 1364
      :::::|||||
401 AsnArgValIleGlnArgLeuArgSer..... 409
      :::::|||||
1365 TTTTAACTTTGAAAAAGTTAAACATACTTTGTATCAGCAATAGCAAAA 1414
      :::::|||||
410 .....GluIleAspHisValIle 416
      :::::|||||
1415 AAGAGAAAGTGACCATTCAAAACTGTTCCGAGAGGCTGATTGCTTAA 1464
      :::::|||||
416 ysgInIleSerAsnLeuGlnGlnSerIleSerAspAlaGlnArgGly 432
      :::::|||||
1465 GAAGTGCCTAATTATAAGCAACTAAAGATGAGAAA..... 1500
      :::::|||||
433 Gln.....AsnAlaLeuLysAspAlaLysAsnLysLeuAsnAs 445
      :::::|||||
1501 .ATGGAAGAAATCATCGCTCAAAATGGCGAGCGATCACCTCAAGACAG 1549
      :::::|||||
445 pLeuGlnAspAlaLeuGln..... 452
      :::::|||||
1550 TTGATGATCTTATCGCAAAAGGTAACGCAAAAATTACCAAGATGAGCTA 1599
      :::::|||||
453 .....AlaIysGlnAspLeu 457
      :::::|||||
1600 TCAAAAGTTGTGATACTAGTAATGCTCAAAACATAGCAAAAATGTGAC 1649
      :::::|||||
458 ThrArgLeuLeuArgAspTyrGlnGlnIleLeuMetAsnThrLys..... 471
      :::::|||||
1650 AAACAGCTTGATAGTTAATCTCATCTGTAAGCGACTTACCGCTGCTA 1699
      :::::|||||
472 .LeuAlaLeuAspLeuGlnIleAlaIleThrTyrArgThrLeuLeuGlnIlg 488
      :::::|||||
1700 ATGATTTAGAGA...AATGTTATTAGGCTCCCAATCTCAATGTTGATCA 1746
      :::::|||||
488 lueGlnSerAlaGlnMetSerGlyGlnLysAlaProAsnValSerValThrVal 504
      :::::|||||
1747 AGTTATCTCTCTTCAATTTGCTAGGGGATCTGACATTTGAGACTACGG 1796
      :::::|||||
505 SerThrSerHisThrSerIleSerGlyGlySerArgGlyGlyGly 521
      :::::|||||
1797 CTTGCGGCCCTGGCGCGTCTCTCAAGATGG.....AGCT 1831
      :::::|||||
521 yGlyGlyTyrGlySerGlySerSerTyrGlySerGlyGlyGlySerT 538
      :::::|||||
1832 ACGGCTGCGCTGCGGCTGCTAGCCAGCATTTGAGCTAGCGGCTGCGC 1881
      :::::|||||
538 yrgIysSerGlyGlyGlyGlyGlyGlyGlyGlyGlySerTyrGlySerGly 554
      :::::|||||
1882 CTTGCGCAGC...GGTACGCAAGATTGAGCTAGCGGCTGCGCTGCGG 1928
      :::::|||||
555 GLyGlySerTyrGlySerGlyGlySerSerTyrGlySerGlyGlyGly 571
      :::::|||||
1929 ATCTCAGCATTTGAGCTAGCGGCTGCGGCTGCGGCTTCTCAAGATT 1978
      :::::|||||
571 yGlyGlyHisGlySerTyr.....GlySerGlySerSerSerG 584
      :::::|||||
1979 GGAGCTAGCGGCTGCGGCTGCTAGCCAGCATTTGAGCTAGCGGCT 2028
      :::::|||||
584 lyeGlyTyrArgGlyGlySerGlyGlyGlyGlyGlySerSerGlyGly 600
      :::::|||||
2029 CTGCGCCCTGCGCAGCGGTAACCAAGATTGAGCTAGCGGCTGCGGCG 2078

```

```

seq_name: p1r2:C83339
seq_documentation_block:
  hypothetical protein PA2462 [imported] - Pseudomonas aeruginosa (strain PA01)
  C:Species: Pseudomonas aeruginosa
  C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
  C:Accession: C83339
  R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
  adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L.
  .; Lory, S.; Olson, M.V.
  Nature 406, 959-964, 2000
  A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
  A:Reference number: A82950
  A:Accession: C83339
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-5627 <STC>
  A:Cross-references: GB:AE004673; GB:AE004091; NID:g9948501; PIDN:AA05850.1; GSPDB:GN
  A:Experimental source: strain PA01
  C:Genetics:
  A:Gene: PA2462

alignment_scores:
  Quality: 171.50      Length: 725
  Ratio: 0.482        Gaps: 42
  Percent Similarity: 49.103      Percent Identity: 22.483

alignment_block:
  US-09-306-689-12 x C83339      ..
  Align seg 1/1 to: C83339 from: 1 to: 5627

31 TGGAGCTACGGCCCTGCGCCCTGCGAGCGGTTCTCAATTTGAGACTACGG 80
      |||:::|||||
5018 TrrpLysIser.....AsnSerGly.LeuGlyValGlyAlaGlnL 5031
      :::::|||||
81 CTTGCGCTGCGGCTGCTAGCC.....AGCATTTGAGCTAGCGCTGCG 124
      :::::|||||
5031 euAlaIleAlaIleValAlaIleTyrPheThrAlaGlyAlaIleSerAla 5047
      :::::|||||
125 GCCCTGCGACGCGGTAGCCAAAGATTGAGCTAGCGGCTGCGGCTGGA 174
      |||:::|||||
5048 AlaLeuGlnSerMetAlaGlyValGlyAla..... 5057
      :::::|||||
175 TCTCAGCATTTGAGACTACGGCTGCGGCTGCGGCTGCTCTCAAGATTG 224
      |||:::|||||
5058 .....GlySerGlySerMetMetAlaAlaAla.....G 5067
      :::::|||||
225 GAGCTAGCGGCTGCGGCTGCGGCTGCTAGCCAGCATTTGAGACTAGCGG 274
      |||:::|||||
5067 lYserThrAlaMetValGlnAlaGlyThrAla...ValGlyThrAlaAla 5082
      :::::|||||
275 TGCGCCCTGCGACGCGGTAGCCAAAGATTGAGCTAGCGGCTGCGGCT 324
      :::::|||||
5083 .....AlaGlyTrrpAlaAsnAlaAlaGlyThrAla...ValAlaMe 5095
      :::::|||||
325 GGATCTAGCTTTCCAAAAGCTGGGGCAAAAAATTTCTCTCTATTTCC 374
      :::::|||||
5095 tGlyMetAlaSer.....AsnGlyAlaIle 5104
      :::::|||||
375 CCAAAATTACCAATAT.....GATACGAA 399
      :::::|||||
5104 ertThrIleAsnAsnArgGlyAsnLeuGlnLysPalaValLysAspValThr 5120
      :::::|||||
400 CAAAGTAATGTTTACAGATTAGTCAAAAGCGCGCAAGAGTTGGGAT 449

```

5121 SerSeraspAlaLeuArgLytyrValValAlaIaGlyThrThrAlaGlyLe 5137  
5122 .....  
450 T.....GAGCTCAAAAGAGAGACGCAATTAATATTCGAA 484  
5137 ThrThrAlaGlyValTyrAspLysTrpThrSerThrGlnThrGlyThrSerT 5154  
485 CAGCTCAAAACGCTTATAGGACGATTCAAACCGCATATGGCTTA...ACT 531  
5134 hTAlaLeuProfsnThrGlyAlaValAlaProAlaAlaGlyLeuGlyThr 5170  
532 GAGCGTCATTTGTGTATCCGCTCCACAAATTGATTAATTCCTACAGAA 581  
5171 TrpGlnGlyVal..... 5174  
582 AACTTAAGACGCCCAAGCATTAAGTTCTCCGCAAGCATTTGACAAATG 631  
5175 .....GlyGlnPhe.....ThrSerAsnGlnLeuLeuGlnAsn. 5185  
632 CAATTAAGCCAAACCTGATATCTGCATTCATTTATTTAGGCTCA 681  
5186 .....GlyThrSerValLeuLeuAspArg 5193  
682 GTATTGGCTGGAATG...GATTAGATGAGCGCTTACAGATAACAGCAA 728  
5194 AlaLeuGlyGlyLeuGlySerLeuGlyAspAlaLeuGlnAsn..... 5207  
729 CCACATGCTCTTGTCTAAAGCTGGCTTGAGCTTAACAAATTCATTAATTG 778  
5207 ..... 5207  
779 AAAATATGCTAATTCAGTAAACACCTGACGAATTTGGTAGCAAAAT 828  
5208 ..SerLeuAlaAsnAla.....Phe 5213  
829 AGTCAATTTGGTTCAAACTCAAAATATCAAAAGCTTAGGACCTTAG 878  
5214 AlaAlaTyrGlyPheLeu.....IleGly 5222  
879 AGACAACCTCAAAATATTCGCTGAGCTTGATTAAGCTGGCTTGGTTAG 928  
5222 yAspThrThrHis.....GlyValLeuAspAspLysSerLeuGlyLysI 5237  
929 ATGTTATCTCAGGCGCTATATCGGGCGCAACGCTGCACTGTACTTGA 978  
5237 legLyLeuHisAlaLeuMetGlyGlyLeuAlaIaGlyAlaValGlyLys 5253  
979 GATTAATAATGCTTCAACAGCTAAAAAGAGGGGTGGTTTGAATTTGC 1028  
5254 AspPheAlaGlyAla.....LeuAlaAlaGlyValAlaAsnGlyAla 5267  
1029 AAACCAAGTTGTGTGAATATATTCACAAAGCGTTCTTCTTACATTTAG 1078  
5267 a.....LeuValAspSerLeuAlaLysGlnTyrAlaSerLeuProIle 5282  
1079 CCAACAGCTGTGCACAGCTTATCTCACTGGCGCTGGCTGCTTTA 1128  
5282 spAspLysLysGlyLeuLeuIleMetSerSerGlnLeuIleGlyValLeu 5298  
1129 ATTCGCTACGTTCTCTCTGCGATAGCCCATTAAGA..... 1167  
5299 AlaAlaSerThrGlnGlyAspAlaAspAlaLysSerLeuGlnThrGlyAl 5315  
1168 ...TTTGCCGCTATTCGCGATTAATTAATCATGCAAAAGCTTAGACA 1213  
5315 atyrValAlaGlyAsnAlaThrGlnHisAsnTyr.....LeuSerH 5329  
1214 GTTATGCCGAGCGCTTAAAAATTAAGCTATGAGCGA.....GATAT 1257  
5329 IstTrpGlnGlnLysLysArgGlnGlyAlaAspLysLysAspLys 5345  
1258 TTATTAGCAGATATTCAGCGGAGCAAGGACTATTGATGCAATCGTTAC 1307  
.....

5346 GlnLeuCysLys.....ThrGly...IleGlyAlaLysTrpAl 5357  
1308 TGCATTAATACC.....GCATTGGCCGCTATTGCTGGTG 1342  
5357 alIleSerAlaGlnGlnAspValGlyTLeValValGlyValGlyG 5374  
1343 GGTGTCTGCTGCGAGCGATTTA.....ACATTTGAAAAAGTT 1383  
5374 LyIleGlyLeuSerThrAlaGlnTyrAlaValGlyAlaTyrGlnLeuVal 5390  
1384 AAACATTAATCTTGTATCAGATAGCAAAAGCAAAAGATGACCATTTCA 1433  
5391 Lys...AsnTrpArgGlnThrTyrAlaAlaLeuGlnGlnLeuAlaThrSe 5406  
1434 AACCTGTTCCGAGAGCT.....GATTTGCTAAAGAGTGCCTA 1474  
5406 rProGlnPheArgGlnGlnPheGlyAspAsnTyrLeuLysGlyLeuG 5423  
1475 ATTATTAAGCACTAAAGATGAGAAATGCAAGA..... 1509  
5423 LuArgAlaAlaPheLeuThrGlnAlaTyrGlyAspAlaGlyTyrGlnGly 5439  
1510 .....ATCATCGTCAAAATGGCGAGCGATCACCTCAAAAGATTGA 1553  
5440 SerValThrAlaGlyValGlnGlyGlyArgPheAlaAlaGlyLeuValG 5456  
1554 TGAATCTTATGCA...AAAGTAAAGCAAAATTAACCAAGATGAGCTAT 1600  
5456 yAlaLeuThrAlaValLysGlyGlyAlaGlnIleThrAla..... 5469  
1601 CAAAAGTCTTATACATTAATGATTCGCTCAAAACATAGCAAAATGACA 1650  
5470 .....LysLeuProThrAlaAlaLysAsnLeuVal 5479  
1651 AACAGCTTAGATTAATTAATCTCATCTGTAAGTGCATTTACCTGCTCA 1700  
5480 AsnAlaIle.....AlaGlnSerProValSerGlySerLe 5491  
1701 TGAATGAGCAAAATGATTAAGTGCCTCAAACTCAATGTTG.....G 1741  
5491 tSerSerGlnLeuGlyAlaValGlyAspLeuGlyArgLeuGlyGlyG 5508  
1742 ATCAAAAGTTATCTCTCTTCAATTTGCTAGGAGGATTCAGCATTTGAGAC 1791  
5508 LyLysGlyTyrValAspTrpLeuSerHisGlyAlaLysGlnHisIleLeu 5524  
1792 TAGCGCTCGCCCTGGACGCGTTCTCAAGATTTGAGCTAGCGCTGCG 1841  
5525 TyrGlyAspLysProGlySerGlyGlnHisLeuTrp..... 5536  
1842 TCCGGGTGCTTACG.....CAGCATTTGAGCTAGCGGC 1876  
5537 .ProGlyGlnAlaGlyLysThrValPheProGlnAsnTrpSer .AlaAsp 5552  
1877 TGCAGCTCGGCGGCGTACCAAGATTTGAGCTAGCGCTGCTGCGGT 1926  
5553 LysIleValHisGlyValGlyAspIleAlaThrSer .ProSerThrLysT 5569  
1927 GG.....ATCTCAGCATTTGAGCTAGCGGCT 1952  
5569 rPTyrAlaGlnThrGlyThrGlyGlyValTyrThrSerLysGlyAspPro 5585  
1953 GCGCCCTGGACGCGTTCTCA.....AGATTTGA 1981  
5586 AlaLysTrpValAlaTyrGlnValAlaArgAspGlyValArgMetArgValVa 5602  
1982 GCTAGCGCTGCTGCGCGG 2000  
5602 TyrGlnProAlaThrGly 5608  
seq\_name: p1r2:A75564  
seq\_documentation\_block:

conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #ext\_change 17-Mar-2000  
 C:Accession: A75564  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M., Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896  
 A:Accession: A75564  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1467 <WHI>  
 A:Cross-references: GB:AE001870; GB:AE000513; NID:g6457732; PIDN:AAF09665.1; PID:g645773  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR0075  
 A:Map position: 1

alignment\_scores:  
 Quality: 169.00 Length: 838  
 Ratio: 0.426 Gaps: 40  
 Percent Similarity: 47.375 Percent Identity: 22.196

alignment\_block:  
 US-09-306-689-12 x A75564 ..

Align seg 1/1 to: A75564 from: 1 to: 1467

```

43 CTGGCGCCCTGGCAGCGGTTCT...CAAGATTGAGCTACGCGCTGCTCC 89
||||| ||||| ::| ||||| ::| |||||
194 LeuArgspGlyAlaThrAspLeuGlnGlyThrValLysLeuArgAs 210
90 GCGTGGCTCT.....AGCCAGCATTTGGAGCT 115
||||| ||||| ::| |||||
210 pGlyAlaSerAspLeuAlaAspGlyAlaAsnArgAlaAlaAspLys 227
116 ACGCCCTGCGCCCTGGCAGCGGTTAGCCAGATTGAGCTACGCGCTGCT 165
||| ||||| ||||| ::| |||||
227 rGlySerLeuAlaAspGlyAlaGlyThrLeuSerSerGlyValGlyArgLeu 243
166 CCGGGTGATCTCAGATTGAGCTACGCGCTGCGCCCTGGCAGCGGTTTC 215
||||| ||||| ||||| ::| |||||
244 ThrAspGlyThrValLysLeuSerAspGlyLeuArgGlnLeuGlnAla 260
216 TCAAGATTGGAGCTACGCGCTGCTCCG.....GGTGGCTTAGCCAGC 259
::| ||||| ||||| ::| |||||
260 AlaProGlyGlnThrGlnLeuAlaProLeuArgSerGlyAlaGln 276
260 ATTGAGCTACGCGCTGCGCCCTGGCAGC..... 288
||||| ||||| |||||
277 .....LeuArgThrGlySerGlnSerLeuAlaSerGlyLeu 288
289 .....GGTAGCCAAAGATTGAGCTACGCGCTGCTGCTCGC. 321
||||| ||||| ::| |||||
289 GlyGlnLeuAlaAspGlyThrAspGlnLeuAlaGlyAlaLysLysAl 305
322 .....GGTGGATCTAGCTTCCCAAAACTGGGGCAAAAAAATTATCTCT 367
||||| ||||| ::| |||||
305 AsnGlyGlyAlaGlnGlnValAlaAlaGlyAsnLysGlnLeuAlaLag 322
368 ATATTCCCAAAATTCACCAATATGATACGAAACAAGTAATGTTTACAG 417
::| ||||| ||||| ::| |||||
322 LnuProGln.....LeuGlnSerGlyLeuThr 331
418 GATTTACTCAAGCGCGGCGAAGACTGGGATTCAGGTACAAAGAGAAGA 467
::| ||||| ||||| ::| |||||
332 GlnLeuSerSerGlyAlaAspLysLeuAla..... 341
468 ACCCAATATATTGCAACAGCTCAAAACAGTTTAGCGACGATTCAACCG 517
||| ||||| ::| |||||

```

```

342 .....AlaGlyAlaLysSerAlaAsnGlyAlaGlnGln 354
518 CTATTGCG.....TTACGTAGCGT..... 537
::| ||||| |||||
354 AlaAlaGlyGlyAsnLysGlnLeuAlaGlnGlnLeuProLysLeuGlnGly 370
538 GGCATTGTGTTATCCGCTCCCAAAATTGATAAATGCTACAGAAACTAA 587
||||| ||||| |||||
371 GlyLeuThrGlnLeuSerGlnGlyAlaAspLysLeuAlaGlyAlaLys 387
588 AGCAGGCCCAAGCATTTAGTTCTGCGGAAGCATTTGTCAAAATGCAATA 637
::| ||||| ||||| ::| |||||
387 sSer.....AlaAsnGlyGlyAlaGlnGlnValAlaGlyLysLysG 402
638 AAGCCAAACTGTATTATTCGCAATTCACCTATTTAGCGCTGATTTG 687
::| ||||| ||||| ::| |||||
402 LnuAlaAspGlnLeuProGlnLeuGlnSerGlyLeuThrGlnLeuSer 418
688 GCTGAATGATTTAGATGAGCCCTTACAGATTAACAGCAACCAACATGC 737
||||| ||||| |||||
419 GlnGlyAlaAspLysLeuAlaAlaGlyAlaLysSerAlaAsnGlyAla 435
738 T.....CTTGCTAAAGCTGGCTTGAGCTACAAATTCATTA..... 774
::| ||||| ||||| ::| |||||
435 aGlnGlnValAlaAlaGlyAsnLysGlnLeuAlaAspGlnLeuProGln 452
775 .....ATTGAAATATTTGCTAAT 792
452 eugInSerGlyLeuGlyGlnLeuSerGlnGlyAlaAspArgLeuAlaSer 468
793 TCAGTAAAAACACTTTCAGCAATTTGTGAGCAAAATTCAGT..... 831
::| ||||| ||||| ::| |||||
469 GlyAlaLysAspAlaAsnAlaGlyAlaGlnGlnValAlaAlaGlyAsn 485
832 CAATTGTGTTCAAACTACAAAATATCAAA...GGCTTAGGCACTTAG 877
||||| ||||| ||||| ::| |||||
485 sGlnLeuAlaAlaGlnLeuProGlnLeuGlnSerGlyLeuGlnLeu 502
878 GAGCAAAACTCAAAATATATC...GGTGGACTGTATAAGCTGGCTGCT 924
::| ||||| ||||| ::| |||||
502 eugInGlyAlaAspArgLeuAlaSerGlyAlaLysGlyAlaAsnAlaGly 518
925 TTAGATGTTATCTCAGG..... 942
::| ||||| |||||
519 AlaGlnGlnValAlaAlaGlyAsnLysGlnLeuAlaThrGlnLeuProGln 535
943 CTATTATCGGGCGCCACACAGCTGACTTGTACTTGCGAGATAAAATCTT 991
||||| ||||| ||||| ::| |||||
535 nLeuGlnSerGlyLeuGlnLeuSerGlnGlyAlaAspArgLeuAla 552
992 CAACAGCTAAAAAAGTGGTGGCGGTTT..... 1020
||| ||||| |||||
552 eugAlaLysAspAlaAsnAlaGlyAlaGlnGlnValAlaAlaGlyAsn 568
1021 .....GAATTGGCAAAACCAAGTT.....GTTGGTAATAT 1049
::| ||||| ||||| ::| |||||
569 LysGlnLeuAlaThrGlnLeuProGlnLeuGlnSerGlyLeuGlnLe 585
1050 TACCAAAAGCCGTTTCTTACATTTAGCCCAAGCGTTGACAGAGTT 1099
::| ||||| ||||| ::| |||||
585 uSerGlnGly.....AlaAspArgLeuAlaSerGly 596
1100 TATCTTCACTGAGCGCTGCTGCTTATTTAGCTTCTTACTGTTTCTCT 1149
::| ||||| ||||| ::| |||||
596 LAsnAspAlaAsnSerGlyAlaLysGlnLeuAla..... 607
1150 GCGATTAGCCATTAGCATTTGCCGCT.....ATTGCCGATAAAT 1190
::| ||||| ||||| ::| |||||
608 .....AlaGlyAsnLysGlnLeuAlaAspGlnLe 617
1191 TAATCATGCAAAAAGT.....TTAGAGAGTTATGCCGAAGCT 1228
::| ||||| ||||| ::| |||||
617 uProGlnLeuGlnSerGlyLeuGlnLeuSerGlnGlyAlaAspArgL 634

```

```

1229 TTAATAATTAGCTATGAC..... 1248
      :: :: :: |||
634 euAlaSerClYAlaAsnAlaAlaGlyAlaAlaGlnIleAlaAla 650
1249 GGAGATAATTATTAAGCA.....GAAATACAGCGGGGAACAG 1286
      ||||| ::|||
651 GlyAsnAlaGlnLeuAlaThrGlnLeuProGlnLeuGlnSerGlyLeu 667
1287 GACTATGATGATCGGTTACTGCAATTAATACCGATTGGCCGCTATG 1336
      ::||| ::|||
667 nGlnAlaSerAlaGlyAlaAspArgLeuAsnThrGlyAlaAlaGlnLeu 684
1337 CTGATGCTGTGCTGCTGCTGCGAGCGGATTAACATTGAAAAAGTAA 1386
      || |||||
684 laThrGlyThrLeuGlnLeuAsn.....AspAlaValLys 695
1387 CATTAATCTTGATCATGCAATAGCAAAAAAGAAAGTACCATTCAAA 1436
      ::||| ::|||
696 AsnGlnAlaLeuLeuProGlyGlnLeuLysGlnGlyAlaAlaLysVal 712
1437 CTGGTTCGAGAGCGCTGATTTCCTAAAGAACTGCTAATTATTAAGCA 1486
      || |||||
712 n.....SerGlyAlaMetGlyValHisSerGlyThrAsnT 724
1487 CTAAAGATGAGAAATCGAAGAAATCATGTCGCAAAATGCGCGGAGTC 1536
      || |||||
724 hLeuSerAsnAlaLeuProGlnLeuGlnAlaGlyAlaGlyThrGlyAla 740
1537 ACCTCA.....AAGCAAGTTGATCTTATCGCAAAAGTAA 1574
      ||||| ::|||
741 ThrGlyAlaSerGlnLeuAlaGlnGlyAlaAspLysLeuAlaGlyLys 757
1575 CGGCAAAATTAACCAAGATGACCTTCAAAAGTTGTTAACTGTAAT 1624
      ::|||
757 rAlaGlnLeu..... 760
1625 TCCTCAAAACATGCAAAATGTGCAAAACAGCTTAGATAAGTTA..... 1668
      |||||
761 .....SerGlnGlyAlaThrThrLeuSerSerLysLysGln 773
1669 .....ATCATCTGTGAAGTGCATTACCTGCTCAATGA 1703
774 AlaGlyGlnGlySerThrAlaAlaValGlnGlyValGlnLeuSer 790
1704 TTCGGAATGTATTAGTGGCTCCAACTCAATGTTGATCAA..... 1746
      ::|||
790 nGlyAlaThrArgLeuAlaAlaGlyThrAlaGlnLeuSerGlnGlyAla 807
1747 ...AGTTTACTCT..... 1758
807 alThrLeuSerSerLysLysGlnAlaGlnGlnGlyThrThrAlaAla 823
1759 CTTCAATTTGTAGGAGATCTCAGCATTTGAGCTACGCGCTCGCC.... 1804
      ::|||
824 ValGlnGlyAlaGlnGlnLeuGlyGlnGlyAlaAspLysLeuAlaAla 840
1805 .....CTGGCAGCGGTTCTCAAG 1822
840 lYThrAlaGlnLeuSerGlnGlyAlaAsnThrLeuSerSerLysLys 856
1823 .....ATTGAGCTACGCGCGCGCGGCTGCTCAAGCAGCAT 1863
857 GluAlaGlnGlnGlyThrThrAlaAlaValGlnGlyAla.....GlnGln 872
1864 TCGAGCTACGCGCTCGCGCTCGCAGCG..... 1891
872 nGlyGlnGlyAlaAspLysLeuAlaAlaGlyThrAlaGlnLeuSerGln 889
1892 .....GTAGCCAAAGTTGAGCTACG 1912
889 lYAlaThrThrLeuSerSerLysLysGlnAlaGlnGlnGlyThrThr 905

```

```

1913 GCCTGCTCCGGGTGAGATCTCAGCATTTGGAGCTACGCGCTCGCCCTGCGC 1962
      |||||
906 AlaAlaValAlaGlnGlyAlaGlnGlnLeuGlnGlyGlnGlyAlaAspLysLeuAl 922
1963 A.....GCGGTTCTCAAGATTGAGCTACGCGCTCGCGGGTGG 2003
      |||||
922 alAlaGlyThrAlaGlnLeuSerSerGlnGlyAlaThrThr...LeuSerThrL 938
2004 CTTAGCCAGCAATTGGA...GCTAGCGCTCGCGCTCGCAGCAGCTACGC 2050
      |||||
938 yAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 954
2051 AAGATTGAGACT 2062
955 lYsLeuAlaAla 958

seq_name: p1r2:A82615
seq_documentation_block:
surface protein XP1981 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: A82615
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: A82615
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1190 <SIM>
A:Cross-references: GB:AE004017; GB:AE003849; NID:9107083; PIDN:AAF84783.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.D.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kempner, E.L.; Kitejima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Sa, R.G.; Santelli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1981

alignment_scores:
Quality: 168.50 Length: 847
Ratio: 0.428 Gaps: 40
Percent Similarity: 46.517 Percent Identity: 21.842

alignment_block:
us-09-306-689-12 x A82615 ..
Align seg 1/1 to: A82615 from: 1 to: 1190
40 GCGCTCGCGCTCGCAGCGGTTCTCAAGATTGAGCTACGCGCTCGCTCC 89
      |||||
155 GlyLeuValAlaAspLysSerGlyAla.....ThrYSerMetArgLe 168
90 GGGT.....GCGCTAGCCAGCATTTGAGCTACGCGC 121
168 nGlySerIleAlaThrMetAsnGlySerAlaGlyLysAspSerIleAlaI 185
122 TCGCGCTCGCAGCGCTAGCCAA.....GATTGAGCTACGCGCTG 162
      ::|||
185 lGlySerClYGlnGlySerLysThrAspGlyAsnThrSerClYAlaThr 201

```

```

163 CGTCCGGGTGATCTACGATTGGAGCTACGCGCTGCGCCCTGCGCAGC 212
    ||| :||| :|||
202 ValAlaIeInGlyLeuArgSerIleAlaIleGlyThrThrAlaArgSerG 218
213 TTCTCAAGAT...TGAAGCTAGCGCGCTCGCGGTGCGCTAGCAGC 259
    ||| ||| |||
218 nSerGlnAspAlaIleSerIleGlyThr.....GlyAlaSerThrT 232
260 ATTGGAGCTACGCGCGCTGCGCGCGGAGCGGTAGC...CAAGATTGAGC 306
    :||| :||| :|||
232 hrGlyAsnPhenAlaIleAlaIleGlyAsnGlyAlaLeuThrSerIleAla 248
307 TACGCGCTGCGCGCTGCGGTGATCTAGCTTCCCAAAAACCTGGGCAAAAA 356
    ||| :||| :|||
249 AsnGlyIleAlaLeuGlyAlaSerSer..... 257
357 AATTATCCTATATTCCTCCCAAAATACCAATATGATGATGACAGAGTA 406
    :||| :||| :|||
258 .....SerValThrThrArgGlyGly 265
407 ATGCTTACAGGATTAGTCAAGCGCGCGAGAGTTGGGATGGAGTA 456
    :||| :||| :|||
265 AlaIaLeuGlyInGlySerLeuAlaIaThrAlaSerGlyIleThrGly 281
457 CAAAGAGAGAGCAGCAATATATTCAGCAAGCTCAACACAGTTAGGCAC 506
    :||| :||| :|||
282 TyrAspProValThrLysSerThrSerThrLeuSerThrSerMetTrp.. 297
507 GATTCAAACCGCTATTCGCTTACTGAGCGCTGCGATTGTTGATCCGCTC 556
    :||| :||| :|||
298 ....ArgSerThrLeuGlyAlaValSerIleGlyAsnIleThrSerSer 313
557 CACAAATTGATTAATTCGCTACAGAAATCAAGAGCGCAAGCTTTCAGT 606
    :||| :||| :|||
313 hrSerGlnThrArgGlnLeuThrGlyLeuAlaIaGlyArgSerAspThr 329
607 TCTCCGCAAGCATTTGACAAATGCAAAATCAAGCAAACTGTA.... 651
    :||| :||| :|||
330 AspAlaValAsnValAlaGlnLeuLysLeuLeuAlaGlySerValGlyG 346
652 .....TTATCTGGCATTCATTCATTTTTCAGTTCAGTAT 685
    :||| :||| :|||
346 yGlyTrpAsnLeuThrAlaSerGlyAlaAsnSer.....SerAsnValA 361
686 TGGGTGAATGATTTAGATGAGCGCTTACAGATACAGCAACCAACAT 735
    :||| :||| :|||
361 LeuGlyGlyLysSerValAsp.....LeuLysAsnSerAspGlyAsnLeu 375
736 GCTCTTCCTAAAGCTGGCTTGAGCTAACAAATTCATTAATTGAAATAT 785
    :||| :||| :|||
376 LeuIleThrLysThr.....ThrAspSerAsnSpValThrPheAsnLe 390
786 TGCTAATTCAGTAAAA..... 801
    :||| :||| :|||
390 uAlaThrAlaLeuLysValAspSerLeuThrGlyAsnThrAlaMetT 407
802 .....ACACTTGAGCAATTTGCT 819
407 hrThrAspGlyValThrValGlyLysArgValThrLeuAspSerThrGly 423
820 .....GAGCAAAATTAGTCAATT 836
424 LeuValIleAlaGlnGlyProSerValIleSerSerGlyIleAsnAlaI 440
837 TGGTTCAAAACTACAAATATCAAAAGCTTAGGACTTTAGAGACAAAC 886
    :||| :||| :|||
440 agGlyGlnLysIleMetAsnVal...GlyThrGlyThrAlaAspThrAspa 456
887 TCAAAATATTCGTGGACTT.....GATTAAGCTGGCTTGGT 924
    :||| :||| :|||
456 IavaIAsnPhenGlyGlnLeuGlnAlaValSerAspThrAlaSerLysGly 472

```

```

925 TTAGATGTTATC.....TCAGGCGCTATTATCGGCGC 956
    :||| :||| :|||
473 TrpAsnLeuLeuAlaSerGlyThrAsnSerSerAsnValAlaProGly 489
957 AACAGCTGCACCTTACTTACGATTAATAATGCTTCAACAGCTTAAAAAG 1006
    :||| :||| :|||
489 aSerValAspLeuLysAsnThrAspGlyAsnLeuIleThrLysAlaI 506
1007 TGGGTGGCGGT.....TTTGAAATTCGCAAAACCAAGTT..... 1038
    :||| :||| :|||
506 IeGlyIleAsnAspValThrPheAsnLeuAlaThrAlaLeuGlnAlaAsp 522
1039 .....GTTCGTAATATTCACCAAGCGCTTCTTCTTACATTTTACG 1079
    :||| :||| :|||
523 SerLeuThrThrGlyAsnThrAlaMetThrThrAspGlyValThrValG 539
1080 CCAACGTTGTCAGCAGCAGCTTATCTTCACTGGCGCTGCTGCTTAA 1129
    :||| :||| :|||
539 ySerAsnValThrLeuGly.....SerThrGlyLeuValIleThrAspG 554
1130 TTGCTTCTAGCTTCTCTTCCGATTCAGC.....CCA 1161
    :||| :||| :|||
554 LyrProSerValThrSerSerGlyIleSerAlaGlyAsnGlnLysIleThr 570
1162 TTACCATTTGCGCGTATTCGCGAT.....AAATTAATTCATGC 1199
    :||| :||| :|||
571 AsnValAlaIaGlyThrAlaAspThrAspAlaValAsnPhenSerGlnLe 587
1200 AAAAAGTTTACAGAGTTATGCCGACGCTTTAAAAAATTAGCGTATACG 1249
    :||| :||| :|||
587 uGlnAlaValSerSerThrAlaSerLys.....G 597
1250 GAGATTAATTTATTCAGCAATATTCAGCGGGA..... 1281
    :||| :||| :|||
597 LyrTrpAsnLeuLeuAlaSerGlyAlaAsnSerSerAsnValAlaProGly 613
1282 .....ACAGGGAC 1289
614 GluSerValAspLeuLysAsnSerAspGlyAsnLeuLeuIleThrLys 630
1290 TATGATGTCATCGCTTACTGCAATTAATACCGCATTCGGCGCTTTCCTG 1339
    :||| :||| :|||
630 rThrAspSerAsnAspValThrPheAsnLeuAlaThrAlaLeuLysValA 647
1340 GTGGTGTGTCGCTGCTGACCGCATTTAATTCATTTGAAAAAGTT..... 1383
    :||| :||| :|||
647 spSerLeuThrThrGlyAsnThrAlaMetThrThrAspGlyValThrVal 663
1384 .....AACATATATCTGTGCATCAG..... 1404
    :||| :||| :|||
664 GlySerAsnValThrLeuGlySerThrGlyLeuValIleThrAspGlyPr 680
1405 .....AATACCAAAAAGAGAAAGTGCACATTC 1432
    :||| :||| :|||
680 oSerValThrSerSerGlyIleSerAlaGlyAsnGlnLysIleThr.... 695
1433 AAAACTGTTCCGAGAGCGCTGATTTGCTTAAGCAAGCTTAATTTATAAA 1482
    :||| :||| :|||
696 .....AsnValAlaIaGlyThrAlaAspThrAsp 705
1483 GCAACTAAAGATGAGAAATCGAAGAATC..... 1512
    :||| :||| :|||
706 AlaValAsnPhenSerGlnLeuGlnAlaIaValSerSerThrAlaSerLysG 722
1513 .....ATCGGTCAAAAT.....GGCG 1528
    :||| :||| :|||
722 yTrpAsnLeuLeuAlaSerGlyAlaAsnSerSerAsnValAlaProGlyG 739
1529 AGCGATCACCTCAAGCAAGTTGATGAT.....CTATCGCAAAAAGT 1572
    :||| :||| :|||
739 LysSerValAspLeuLysAsnThrAspGlyAsnIleValIleSerLysGlu 755
1573 AACGCAAAATTTACCAAGATGAGCTATCAAAAGTTGTTGATACTATGA 1622

```



```

      756 SerGly.....SerAsnAspValLeuPheAsnLeuSerSerSerLeu 770
      1623 ATG..... 1626
      770 SLeuAspLysLeuThrValGlyAspThrValMetThrThrAsnGlyVal 787
      1627 .....CTCAACAT 1635
      787 hrValGlySerGlyValThrLeuGlySerMetGlyLeuValIleThrAsp 803
      1636 AACAAATATGTGACA.....AACAGCTAAGATTAATCTC 1673
      804 GlyProSerValThrSerSerGlyIleAsnAlaGlySerGlyLysIle 820
      1674 ATCTGTAAAGTCATTACCTCGCTAAAGATTCGAGAAATGATTAGTGG 1723
      820 rAsnValAlaAlaGlyThrAlaAspThrAspAlaValAlaAsnLeuSerGln 837
      1724 CTCCACTTCATGTTGGATCAAGTTATCTTCTCTCAATTGCTAG 1773
      837 euAsnThrAlaMetAlaGlySerGlyAlaLysSerValHisTyr..... 851
      1774 GGATCTCAGCATTTGAGCTACGCGCTGCGCCCTGCGACGCGTTCTCAGA 1823
      852 ..TyrSerThrTyrAspGlyGlyThrGlnGlyGlyAsnTyrAsnGlyAs 867
      1824 TTGAGACACGCGCTCGCTCGCGGTGCTCTAGCCAGCATTTGAGCTAG 1873
      867 p.GlyAlaThrGlyThrArgSerIleAlaVal..GlyValGlyThrLeu 882
      1874 GCCTGCGCGCTGCGACGCGGTAGCCAGATTTGAGCTACGCGCTCGCG 1923
      883 AlaSerAla.....GluGlyAlaThrAla..... 890
      1924 GGTGATCTCAGCATTTGAGCTACGCGCTGCGCGCTGCGACGCGTTCTCA 1973
      891 .....ValGlySerGlyAlaAlaAlaAlaSerGlyLys..... 900
      1974 AGATTGAGCTACGCGCTCGCGGTGCTCTAGCCAGCATTTGAGCT 2023
      901 .....GlySerThrAlaIleGlyArgAsnAlaValAlaSerAlaAspGly 915
      2024 ACGGCTCGCGCTGCGACGCGGTAGCCAGATTTGAGCT 2062
      916 SerVal...AlaLeuGlyAspGlyAlaLysAspGlyAla 927

seq_name: p1r2:T14961

seq_documentation_block:
hypothetical protein Y1044 - Yersinia pestis plasmid pMT1
C:Species: Yersinia pestis
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 22-Oct-1999
C:Accession: T14961; T14657
R:Rindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.
Infect. Immun. 66, 5731-5742, 1998
A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis K1M5 plasmid
A:Reference number: Z18268; MUID:99043898
A:Accession: T14961
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1525 <IN>
A:Cross-references: EMBL:AF074611; NID:g3883003; PID:g3883044; PIDN:AAC82704.1
R:Hu, P.; Elliott, J.; McCreedy, P.; Skovronski, E.; Garnes, J.; Kobayashi, A.; Cariano,
submitted to the EMBL Data Library, March 1998
A:Description: Structural organization of virulence determinants in three Yersinia pestis
A:Reference number: Z18168
A:Accession: T14657
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1525 <RUP>
A:Cross-references: EMBL:AF053947; NID:g2996286; PID:g2996290; PIDN:AAC13170.1
C:Genetics:

```

```

A:Gene: Y1044
A:Genome: plasmid pMT1

alignment_scores:
  Quality: 166.50      Length: 883
  Ratio: 0.386
  Percent Similarity: 48.811      Percent Identity: 21.065

alignment_block:
US-09-306-689-12 x T14961 ..
Align seg 1/1 to: T14961 from: 1 to: 1525

23 CTCAGCATTTGAGCTACGCGCTGCGCGCTGCGACGCGTTCTCAAGATTGG 72
   |||||
513 ILeSerLeuGlyLysValAlaAlaMetValAlaPhe..... 525
73 AGCTACGCGCTGCGGTGCTCTAGCCAGCATTTGAGCTAGCGGC 121
526 .....LysLeuMetArgAlaGlyIleAlaGlyValIleGlyThrAlaG 540
122 .....TGCCTCGCGACGCGTA 139
540 IyGlnMetValAsnThrPheThrLysMetSerThrValLeuGlnAlaPro 556
140 GCCAAGATTGAGCTACGCGCTGCGGTGAGTCTCAGCATTT..... 184
557 PheAsnLeuGlyAlaThrAlaValThrArgPheAsnArgAlaAlaArgme 573
184 ..... 184
573 tGlyLeuAlaProIleProSerLeuIlePheAlaIleArgGlyAlaIleT 590
185 .....GGAGCTACGCGC.....TGCCTCGCGACGCGTTCTC 217
590 hngIyLeuGlnGlyAlaPheAlaGlyLeuThrAlaPheIleAlaAlaAsn 606
218 AAGATTGAGCTACGCGCTGCGTC..CGGTGCTCTAGCCAGCATTT.. 262
607 ProIleGlyAlaAlaPheThrValAlaThrValAlaValAlaGlyLeuI 623
263 .....GGAGCTACGCGCGCTGCGCGCGCGCGCG 290
623 eThrTyrMetThrMetLeuArgSerGluThrSerLysValValAspGluI 640
291 TAGCCAGATT.....GGAGCTACGCGCGCTGCGGTGAGTCTAGCT 334
640 IeArgLysIleProGluAlaMetThrAlaAlaLysArgAlaGlnMetAla 656
335 TCCCA..... 339
657 AlaArgAlaAlaGlyLeuGlnLysGlnIleGlnArgAspGlnAlaIle 673
340 .AAACTGGGCAAAATAATATCTCTATATTCGCCAAATATTACAT 388
673 uysThrGlyGluSer.....ValAsnTyr...T 682
389 ATGATCTGTAACAAGCTAATGTTTACAGATTATGCAAGCGCGCGAA 438
682 ySerThrAlaAlaGly.....ProValAlaValLysGluSerLys 695
439 GAGTTGGGATTTGAGTACAAAGAGAAGACGCAATATATTGCAACAGC 488
696 GluVal.....ValGluAlaArgLeuLysLysAsnGlnGluLys 709
489 TCAAACCACTTACGACGATTCAAACCGCTATTTGCTTACAGACGCG 538
709 rGluArgThrThrGlyThrMetAlaLeuGlyAspGlyAlaValAlaLys 726
539 GCATTGTGTTATCCGCTCCCAAAATGATTAATTTGCTACAGAAACTAA 588
726 rGluAlaLysGluAla.....AlaGluSerGlnIleGluLysIleArg 740

```

[illegible]

1309 palaglutrhrllysalaVallyserIlePheSerMetGlnAsnLeuS 1326

2018 ..GGAGCTACGGCTGGCCCTGGCAGCGGTAGCCAGATTGGAGCT 2062

1326 erSeValThrGlyAlaLeuSerAlaAlaPheAlaMetLeuGlyAla 1341

seq\_name: p1r2:I53169

seq\_documentation\_block:

Cytokeratin 2 - human

C:Species: Homo sapiens (man)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 13-Aug-1999

C:Accession: I53169

R:Collin, C.; Ohayoun, J.

Differentiation 51, 137-148, 1992

A:Title: Suprabasal marker proteins distinguishing keratinizing squamous epithelia: Cyt

A:Reference number: I53169; MUID:93114504

A:Accession: I53169

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-638 <RES>

A:Cross-references: GB:M9063; NID:G181389; PIDN:AAA35746.1; PID:G181390

C:Superfamily: cytoskeletal keratin

alignment\_scores:

Quality: 166.00 Length: 734

Ratio: 0.505 Gaps: 37

Percent Similarity: 44.823 Percent Identity: 23.161

alignment\_block:

US-09-306-689-12 x I53169 ..

Align seg 1/1 to: I53169 from: 1 to: 638

40 GGCCTGCGCCCTGGCAGCGTTCTCAAGATTGGAGCTAGCGCTGCT... 87  
 45 GlyPheArgSerGlyAlaGly.....SerPheGlySerArgSer 57  
 88 .....CCGGGTG 94  
 57 rleuTyraAsnleuGlySerAsnLysSerIleSerIleSerValAlaAlaG 74  
 95 GCTTACGACGACATTGGAGCTACGCGCTGCGCTGGCAGCGGTACCA 144  
 74 LysSerSerArgAlaGlyPheGlyPheGlyGlyArgSerSerGlyPhe 90  
 145 GATTGGAGCTAGCGCGCTGCGGTGGATGATCAAGATTGGAGCTACGG 194  
 91 AlaGlyGlyTyrcGlyGlyPheGlyGlySerTyrcGlyGlyPheGly 107  
 195 CCGTGGCCCTGGCAGCGTTCTCAAGATTGGAGCTAGCGCGCTGCGCG 244  
 107 yGlyGlyArgGlyValGlySerGlyPheGlyGlyAlaGly...GlyPheG 123  
 245 GTGGCTTACGACGACATTGGAGCTACGCGCTGCGCTGGCAGCGGTAC 294  
 123 LysIValAlaGlyGlyPheGlyGlyProGlyValPheGlyGlyProGlySer 139  
 295 CAAGATTGGAGCTAGCGCGCTGCGGTGGATGATCAAGTTGCCAAAAAC 344  
 140 PheGlyGlyProGlyPheGlyPheGlyProGlyGly.....PhePro...G 153  
 345 TGGGGCAAAAAAATTATC.....CTCTATTATCCCAAAATT 382  
 153 yGlyIleGlnGluValIleValAsnGlnSerLeuLeuGlnProLeuAsnV 170  
 383 ACCAATATATATCTAGCAACAAGTAATGTTTACAGGATTAGTCAAGCG 432  
 170 aGluIleAsnProGlnIleGly..... 177  
 433 GCCGAAGATTGGGATTGAGTACAAAGAAAGAAAGCAATTAATATTTC 482

178 .....GlnValLysAlaGlnGluArgGluGlnIleLeu 188  
 483 AACAGCTCAAAACAGTTTAGCAGC...ATTCAAAACCGTATTGCTTAA 529  
 188 srIleuAsnAsnLysPheAlaSerPheIleAsnLysValArgPheLeuG 205  
 530 CTGAGCGGTGACATTGCTTATCCGCTCCACAAATTATTAATTGCTACAG 579  
 205 LueGlnAsnLysValLeuGluThrLysTrp.....GluLeuLeuGln 219  
 580 AAAACTAAAGCAGCCAGCATTAGGTCTGCC.....GAAG 617  
 220 GlnGlnThrThrGlySerGlyProSerSerLeuGluProCysPheGly 236  
 618 CATTTGACAAATGCAAAATTAAGCCAAACCTGATTATGCTGCAAT 667  
 236 rTyrlleSerPheLeuCysLysGlnLeuAsnSerLeuLeuGlyGluArg 253  
 668 CTATTTTAGGCTCAGTATTGGCTGCAATG.....GATTTA 702  
 253 LysnLeuGlnGlyGluLeuLysSerMetGlnAsnLeuValGluAsn 269  
 703 GATTAGCGCTTACAGATATACAGCAACCAACATGCTCTGCTTAA... 747  
 270 LysLysLysTyrcGluAsnGluLeuLysArgThrAlaAlaGluAsn 286  
 748 ....GCTGCTTGGAGCTAACAATTCATTAAATTGAAATTTGCTAATT 793  
 286 upheValGlyLeuLys..... 291  
 794 CAGTAAACACCTTGACGATTTGGTACCAATTAATGCTAATTTGCTTA 843  
 292 ....LysAspValAsp.....AlaAlaPheMetAsn 300  
 844 AAACCTCAAAATATCAAGCTTAGGAGCTTAGAGCAACCACTCAAAA 893  
 301 LysValGluLeuGlnAlaLysValAspSerLeuThrAspIleValSer 317  
 894 TATCGGTGACCTTGATTAAGCTGCGCTGTTAGATTATCTACAGGC 943  
 317 eleuArgThrLeuTyrcGluMetGluLeu.....SerGlnM 329  
 944 TATTATCGGGCGCA.....ACAGCTGCACTGTTACTTCCAGATAAAT 987  
 329 etGlnSerIleAsnSerAspThrSerValValLeuSerMetAsnAsn 345  
 988 GCTTCAACAGCTAAAAAAGTGGGTGCGGTTTGAATTGGCAACCAAT 1037  
 346 ArgCysLeuAsp..... 349  
 1038 TGTGTGTAATATTACCAAGCCGTTCTCTTACATT.....TTAGCC 1081  
 350 ..LeuGlySerIleIleAlaGluValArgThrGlnTyrcGluIleAla 366  
 1082 AACGTGTTGCACAGGTTTATCTCAACTGCGCTGCTGCTTAAAT 1131  
 366 lArg.....SerLysSerGlnAlaGluAlaLeuTyrc 376  
 1132 GCTTCTACTGTTTCTCTTGCATTAGCCATTAGCCGATTTGCCGATTC 1181  
 377 GlnThrLysLeuGly..... 381  
 1182 CGATTAATTAATCATGCAAAAAAGTTTAGAGATTATGCCGAACGTTTA 1231  
 382 .....GluLeuGlnThrAla..... 387  
 1232 AAAAATTAGGCTATACGAGATTAATTTATAGCAAAATATACGGGGA 1281  
 388 .....GlyArgHisGlyAspAspLeu.....ArgAsn 396  
 1282 ACAGGAGCTATTGATGATCGGTTACTGCAATTATACCGCATTTGGCCGC 1331  
 111



```

599 CATTAGCTTCGCCGAAGCATGTGTAACAAAT..... 636
    |||  ::  :::::|||||  ||
211 allenglthrlystrpgluleuenglmetasnvalglythrag 227
    ::  ::  :::::|||||  ::
637 .....AAGCCAAACTGTATTATCGCATTCATATTATAGGCTC 680
    ::  ::  :::::|||||  ::
228 Proileasnleugluproilephegnglytyrileaspsertleulsar 244
    |||  |||  ::  :::::|||||  ::
244 gtyrleuaspglyleuthralaglthrserglnasner..... 258
731 AACATCCTCTTGCTTAAGCTGGCTGAGTACAAAT.....TCATTA 774
259 .....GluLeuasnmetGlnaspleu 266
775 ATGTGAAATATGCTAATTCAGTAAACACTGACAAATTTGGTAGCA 824
    ::|||  ::  |||  |||||  :::::
267 ValGluAspTyrLysLys.....LysTyrGluAspGluLeuLysar 281
825 AATT.....AGTCAATTTGGCTCAAAACTACAAATATCAAGGCT 865
    ::  ::|||  ::  :::::|||||  ::
281 gthralaalaglunasphvalthrleuylsaspvalaspsna 298
866 TAGGACTTAGGAGCAAACTCAAAATATCGCTGAGCTGTAAAGCT 915
298 lathyrlleuylsvalGluLeuGlnser.....LysValAspleu 312
916 GGCCTTGTTAGATGTATCTCAGGCTATTATTCGGCGCACAGCTGC 965
    ::  ::|||  ::  :::::|||||  ::
313 AsnglGluLeuLysLys..... 320
966 ACTTGACTTGACAGATAAAATGCTCAACAGCTAAAGTGGCGG 1015
    |||||  |||  ::  |||
321 ...ValleuTyrAspAlaGluLeu..... 328
1016 GTTTCAATTTGGCAACCAAGTTGTGTAATATACCAAGCCGTTCT 1065
    |||||  ::  ::|||  ::
329 .....GlnLeuHisGlnSerValThrAsp.....Thr 337
1066 TCTTACATTTAGCCCAACGCTGTGGAGCAGTTTATCTTCACTGAGGC 1115
    ::  |||||  ::  ::  ::|||  ::
338 AsnValIleLeuSerMetAspAsnSerArgAsnLeuAspleuAspserI 354
1116 TGTGCTGCTTAAATGCTTCTACTGTTCTGCGAATTAGCCCATAG 1165
    ::|||  ::  |||  ::  ::|||  ::
354 eillealaglualylsaglntyrGluGluIleala.....Glna 368
1166 CATTTGCCGCTATTCGCATTAATTAATCATGCAAAAGTTAGAGAGT 1215
    |||||  ::  |||  ::|||  ::
368 rgsertylsGluGluAlaGluAlaLeuThrHisserLys..... 380
1216 TATGCCAGACGCTTTAAATAATAGCTATGACGAGATAATTATTAGC 1265
    |||  |||  ::|||  |||||  ::|||
381 TyrGluGluLeuGlnValThrValGlyArgHisGlyAspserLeu.... 395
1266 AGAATATCAGCGGGGAACAGGACATATGATGCATCGTTACTGCAATTA 1315
    |||  ::|||  ::|||  ::
396 .....LysGluLeuLysIleGluIleSerGluLeuA 406
1316 ATACCCATTTGGCCGCTATTCGTGTGTGTCTGCTGACGCCGAT 1365
    ||  ::|||  ::|||
406 snargValIleGlnArgLeuGlnGly..... 414
1366 TTAACATTTGAAAAAGTAAACATATCTTGATCAGCAATAGCAAAA 1415
    |||||  ::|||  ::|||
415 .....GluIleAlaHisValLysGly 421
1416 AGAGAAAGTGACCATTCAAAAGCTGTTCCGAGAGGCTGATTTTGTAAAG 1465
    ::|||  ::|||  ::|||  ::
421 scInCysLysAsnValGlnAspAlaIleAlaAspAlaGluGlnArgGly 438

```

```

1466 AAGTGCCTAATTATTAAGCAACTAAGATGAGAA..... 1500
    ||  |||  |||||  ::
438 lu.....HisAlaLeuLysAspAlaArgsnLysLeuAsnAsp 450
1501 ATCGAAGAAATCATTCGCTCAAAAATGGCGAGGATCACTCAAGCAAGT 1550
    ::|||  ::|||  ::|||
451 LeuGluGluAlaLeuGln..... 457
1551 TGATGATCTTATCGCAAAAGGTACGGCAAAATTAACCAAGATGAGCTAT 1600
    ::|||  ::|||  ::|||
458 .....AlaLysGluAspleuA 463
1601 CAAAGTGTGTGATTAATGATGATTCGTCAAACATGCAAAATGAGACA 1650
    ::|||  ::|||  ::|||  ::|||
463 lathrleuLeuArgAspTyrGlnGluLeu.....MetAsnValLys 476
1651 AACAGCTTAGAT.....AGTTAATCTCATCTGT 1679
    ::|||  ::|||  ::|||
477 LeuAlaLeuAspValGluIleAlaThrTyrArgLysLeuLeuGluGly 493
1680 AAGTGCATTTACCTGCTAATGATTCGAGAAATGATTAATGCTCAAA 1729
    ::|||  ::|||  ::|||  ::|||
493 uGluCysArgMetSerGlyAspLeuSerSerValThrValSerValT 510
1730 CTTCATGTTGATCAAGTTTATCTCTTCAATTGCTAGGGAGATCT 1779
    |||||  ::|||  ::|||  ::|||
510 hrser.....SerThrIleSerSerAsnValAlaSerLysAla... 522
1780 CAGCATTTGAGCTACGACGCTCGCCCTGCGAGGCTTCAACATTTGAG 1829
    ::|||  ::|||  ::|||  ::|||
523 .....AlaPheGlySerGlyArgGlySerSer.....Se 534
1830 CTAGGCTTCGCTCGGCTGCTTACGACATTCGACCTGCGCTTCG 1879
    |||  ::|||  ::|||  ::|||
534 rglGlyGlyTyrSerSerGlySerSerValThrValSerValT 551
1880 GCCCTGCGAGCGGTACGCAAGATGAGCTACGCGCTGCGCTCGGCTGA 1929
    ::  |||||  ::  ::  |||  |||||
551 lserGlySerArgGlyGlySerGlyGlyGlySerIleSerGlyGly 567
1930 TCTCAGCATTTGAGCTACGCGCTCGCCCTGCGAGGCTTCAACATTTG 1979
    ::|||  ::|||  ::|||  ::|||
568 .....GlyTyrGlySerGlyGlyGlySerGlyArgGly 580
1980 GAGCTACGCGCTCGCTCGGCTGCT..... 2007
    |||  |||  |||||  ::|||
580 yserGlyGlyGlySerGlyGlySerIleSerGlyGlyGlyTyrGlys 597
2008 .....AGCCAGCATTTGAGCTACGCGCTCGCCCTGCGAGGCTTAC 2049
    ::|||  ::|||  ::|||  ::|||
597 erGlyGlyLysLysHisSerSerGlyGlyGlySerArgGlySer 613
2050 CAAGATTTGAGCTACGCGCTCGCTCGGCTGATTC 2085
    ::  ::|||  ::|||  ::|||
614 SerGlyGlyGlyTyrGlySerGlyGlyGlySer 625
seq_name: pir2:E75590
seq_documentation_block:
methyI-accepting chemotaxis-related protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75590
R:White, O.; Elsen, J.A.; Heidelberger, J.F.; Hickey, E.K.; Peterson, R.J.;
M.; Shen, M.; Yamathavan, J.D.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75590
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-756 <WHI>

```

A:Cross-references: GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF12432.1; PID:g646072  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DRA0353  
 A:Map position: 2

alignment\_scores:  
 Quality: 166.00 Length: 488  
 Ratio: 0.675 Gaps: 25  
 Percent Similarity: 50.410 Percent Identity: 23.361

alignment\_block:  
 US-09-306-689-12 x E75590 ..

Align seg 1/1 to: E75590 from: 1 to: 756

```

550 TCCGCTCCCAAAATGATTAATGCTACAGAAACTAAAGCGCCAAAGC 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
25 SerLeuProArGAlaProArGTrpLeAspArgLeuSerValAlaGlnLy 41
600 ATTGAGTCTCCCGAAAGC.....ATTGACAAATGCAA 634
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41 sLeuGlyLeuAlaGlyAlaLeuPhePheValProLeuValValAsnValT 58
635 ATAAAGCCAAACTGTATTA.....TCTGCATTT 663
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
58 hrAlaAlaLeuSerThrLeuMetArGlyLeValGlnProValAlaGlyAla 74
664 CAATCTATTTAGGCTCAGTATGCTGGAATGCAATGATTAAGAGCCCTT 713
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
75 GluGlnIleMetProPheValAsnAlaThrThrAspLeuGlnArgAlaLe 91
714 ACAGAAAT..... 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
91 uAspAspLeuThrLeuAlaAlaSerAlaGlnThrSerProAlaAlaAla 108
721 ..AACGACCAACACATGCTCTTGGCTTAAGCTGCTGAGCTAACAAT 768
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
108 rGlnArGAlaGlnGlnAlaLeuThrGlnLeuGlyAlaLeuLeuAspSer 124
769 TCATTAATGAAAATATTGCTAATTCAGTAAACACTGACGAAATTTGG 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 .....GluAlaSerThrAsnProIleSerProLeuProGlnIleAr 138
819 TGAGCAATTAAGTCAA.....TTGGTCAAAACATC 850
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
138 gArGlnLeuArGlnThrLeuGlnPheAlaAsnPheArGAlaGlnValG 155
851 AAAATATCAAGGCTTAGGACATTTAGGAGACAAA..... 885
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
155 LuArGlyThrLeuAlaLeuProProSerGlyAspArgValSerSerTyrGly 171
886 .....CTCAAAAATATC..... 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 ThrValAsnSerAlaValAlaGlnAsnValLeuLysAsnIleThrGlnAla 188
898 .....GGTGA.....CTGATAAG 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
188 rSerLeuAspLeuGlnGlyGlnIleHisThrMetTyrLeuLeuGlnArG 205
914 CTGGCCGTGTTAGATGTTATCTCAGGCTATTTGCGGCGCAACAGCT 963
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
221 AlaGlnLeuLeuAsnAspArgLeuArGArgGlyAlaProAsnAlaSerAl 237
205 lAlaLeuGly..AspGlyLeuMetGlyLeuArGSerAsnLeuValGln 220
964 GCACCTTGACTTGCAATAA.....AATGCTTCAAC 995
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
221 AlaGlnLeuLeuAsnAspArgLeuArGArgGlyAlaProAsnAlaSerAl 237
996 AGCTAA...AAAGGGGTGGGGGTTTGAATGGCAACCAAGTTGTG 1042
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
237 aAspArgThrArGLeuSerAlaLeuValThrHisMetSerThrGlnValG 254

```

```

1043 GTAATATTACCAAGCCGTTCTTCTTACATTTTAGCCCAA...CGTGT 1089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
254 LysGlnThrGlnGlnAlaLeuAsnAspAlaValArGArGProGlnGln 270
1090 GCACAGGTTTATCTTCAACTGGCGCTGTGGCTCTTAAATGCTTGTAC 1139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
271 AlaArGAlaLeuSerSerGlu.....MetGlnGlnLeuSerAsnMe 284
1140 TGTTCCTCTGCGATTAAGCCATTA.....GCATTTGCCGTTATGCGG 1183
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
284 tMetSerLeuValLeuGlnAsnLeuArGAlaAlaAlaSerGlyAlaGly 301
1184 ATAAATTTAATCATGCA.....AAAGTTTAGAGT..... 1215
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
301 LysThrSerGlnAlaLeuIleAspArgAsnLeuGlnLysIleAspArg 317
1216 ..TATGCCAGACGCTTTAAAAATTAGGCTATGACGAGATTAATTATT 1262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 LeuTyrAlaGlnThrSerAlaAlaLeuSerGlnGlnLysAsnPheG1 334
1263 AGCAGAATATCAGCGGGGAACAGGAGCTATGATGATCGGTTACTGCAA 1312
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
334 nAspGlnTyrValArg..LysGlyIleValAsnGlyLeuIleThrLeu 350
1313 TTAATACCGCATTTGCCGCT..... 1332
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
350 euIleThrAlaLeuAlaThrTrpLeuLeuValTyrIleLeuGlnArgIle 366
1333 .....ATTGCTGGTGGTGTCTGCTGCTGCGACCGCATTTAATTTGA 1376
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
367 ValArGlnLeuArGlnLeuThrThrAsnAlaGlnAlaLeuThrAspG1 383
1377 AAAAGTTAAACATAATCTTCTCATCAGCAATAGCAAAAAGAGAAGTGA 1426
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
383 yAsnLeuAsnIleHisLeuProValThrThrSer..AspGlnLeuGly 399
1427 CCATTCAAACCTGGTTCGAGAGCGCTGATTTTGCTAAGAAGTCCCTAAT 1476
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
399 alleuThrSerSerPheAsnGlnAla.....AlaAlaGlnLeu..... 411
1477 TATTAAGCACTAAAGATGAGAAAATCGAAGAAATCATCGTCAAAATGG 1526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 .....ArgThrAsnAlaGlnArGValGlnGln..... 420
1527 CGAGCGATCACCTCAAGCAAGTGTGATGATCTT..... 1560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 .GluArGlnLeuSerGlnArGlnLeuGlnAspHisIleGlyGlnPheLeu 437
1561 .....ATGCAAAAGGTAAAC.....GGCAAAATT 1584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
437 spValThrMetAspIleAlaGlnGlyAspLeuThrLysArgGlyValVal 453
1585 ACCCAAGATGAGCTATCAAAAAGTTGTGATACATGATTAATGCTCAACA 1634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
454 ThrGlnAspIleLeuGlnLysAsnValValAspSerIleAsnLeuMet.... 468
1635 TAGCAAAAATGTGACAAACAGCTTAGATTAAGTAAATCATCTAAGTGG 1684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
469 .....ThrAspGlnLeuAlaAlaThrThrLeuGlnGlnValGln 481
1685 CATTTACCTCGTCT 1698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 lAlaLeuSerAlaSer 485

```

seq\_name: plf2:A56143

seq\_documentation\_block:

surface-array protein homolog sapA2 - Campylobacter fetus (strain 82-40 LP3)  
 C:Species: Campylobacter fetus  
 C:Date: 03-Nov-1995 #sequence\_revision 03-Nov-1995 #text\_change 08-Oct-1999  
 C:Accession: A56143  
 R:Dworkin, J.; Tummuuru, M.K.R.; Blaser, M.J.  
 J. Bacteriol. 177, 1734-1741, 1995

A:Title: A lipopolysaccharide-binding domain of the Campylobacter fetus S-layer protein  
 A:Reference number: A56143; MUID:95204338  
 A:Accession: A56143  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1109 <DMC>  
 A:Cross-references: GB:S57860; NID:913763; PIDN:AB33871.1; PID:913764  
 A>Note: this gene appeared to be silent in strain 82-40 LP3  
 C:Genetics:  
 A:Gene: sapA2

alignment\_scores:  
 Quality: 165.00 Length: 672  
 Ratio: 0.493 Gaps: 34  
 Percent Similarity: 49.851 Percent Identity: 22.917

alignment\_block:  
 US-09-306-689-12 x A56143 ..

Align seg 1/1 to: A56143 from: 1 to: 1109

```

346 GGGGCAAAAAAATTATCTCTATATTTCCCAAAATTACCAATATGATAC 395
      ||| |||
      ||| |||
23 GlysnlystyrtpleuAsptyrAlasnAsnSerleuGlyValse 39
396 TGAACAAGTAATGCTTACAGATTAGTCAAAAGCCGCGAAGATTGG 445
      ||| |||
      ||| |||
39 rSerleuAlasnIleMetleuAspSerProGlyAlaAlaIleAspPheg 56
446 GGATTGAGTACAAAGAGAACGCAATATATTCAGACGCTCAAAACC 495
      ||| |||
      ||| |||
56 lAspSerleuAlaGlyAsnIleLysAspPheValThrLysIleTyr 72
496 AGTTAGGACAGATTCAAACCGCTATTGGCTTAAGCGGTGCATTTG 545
      ||| |||
      ||| |||
73 SerIle.....AlaLeuGlyAsnThrser..... 80
546 GTTATCCGCTCCACAAATGTATATGCTACAGAAACCTRAAGCA.... 591
      ||| |||
      ||| |||
81 .....AspValAspGlyIleAsnTyrTrpThrLysAlaIleT 93
592 .....GGCCAAGCATTAGCTTCGCCGAAGCATTTGTAACAATGCA 633
      ||| |||
      ||| |||
93 hArgIleGlyGlyIlePheThrAspSerLysGlyAsnValIleSerValAla 109
634 AATTAACCAAACTGTATTATCTGCATTCATCAATCTATTTAGGCTCAGT 683
      ||| |||
      ||| |||
110 SerleuSerLysGlyAspLeuIleGly.....AlaMetIleAsnSerI 124
684 ATTGGCTGGAAATGATTTACAT.....GAGGCCCTTACAGA 718
      ||| |||
      ||| |||
124 tValAsnGlyGlySerAlaGlySerLysAlaIlePheGlyAlaLysAla 141
719 ATTAAGACACCAACATGCTTCCTTAAGCTGGCTTGGAGCTA..... 762
      ||| |||
      ||| |||
141 lAlaIleSerAspTyrPheAlaIleAspAlaThrleuGlyLysAspIleSerGly 157
763 .....ACAAATCATTAATTAATAATTTGCTAATTCAGT 797
      ||| |||
      ||| |||
158 LeuAspIleGlyThrThrSerLysLeuIleSerGlyIle.....AsnSerAl 173
798 AAAAACCCTTGACGAATTTGGTACCAATTAGTCAATTTGGTTCAAAAC 847
      ||| |||
      ||| |||
173 aSerAspLeuAspLysValLysSerGlyIle..... 183
848 TACAAAATATCAAGGCTTAGGACTTAGAGACAAACTCAAAATATC 897
      ||| |||
      ||| |||
184 .....AspGlyLeuLysGlu... 188
898 GGTGACTGTATAAGCTGGCTT..... 921
      ||| |||
      ||| |||
189 .....SerIleAspIleAlaGlyLeuAsnLysIleAlaIleuThrThrGluAs 204
```

```

922 .....GGTTAGATGTTATCTCAGGCTATTAT 949
204 nasPThrIleThrGlyThrGlyGlyAspLeuIleSerGlyValAlaG 221
950 CGGGCGCAACAGCTGCACCTTGACTTGACATTAATAATGCTTCAACAGCT 999
      ||| |||
      ||| |||
221 lYThrAla.....AlaGlySerThrLeuAsnProGly 231
1000 AAAAAGTG.....GGTGGGGT.....TTGAAATGGC 1028
      ||| |||
      ||| |||
232 AspLysIleAspGlyGlyAlaGlyAsnAspValLeuLysValAspLeuL 248
1029 AAACCAAGTTGTTGCT.....AATATTACCA 1054
      ||| |||
      ||| |||
248 sAsnAsnPheLysGlyLeuLysAspAspGlyTyrIleLysAsnIleGlu 265
1055 AA.....GCCGTTCTCTTACATTTAGCCCAACGT 1086
      ||| |||
      ||| |||
265 yLsLeuSerleuThrAsnSerSerValSerAsnArgThrPheAspAlaLys 281
1087 GTTGCAACAGCTTATCTTCAACTGGCGCTGTGCTCTTAAATGCTTC 1136
      ||| |||
      ||| |||
282 GlyIleAspGlyLeuGlnThr.....ValAlaIle 291
1137 TACTGTTCTCTCCGATTAGCCCATTTAGCATTGGCCGTAATTCGCAT 1185
      ||| |||
      ||| |||
291 uSerGlyGlyLysGlyIleSerValThrAsnLeuAlaAsnIleValAsp 308
1186 ..AATTTAATCATGCAAA.....AGTTAGAGAT...TAT 1218
      ||| |||
      ||| |||
308 alGluValAsnGlyPheLysGlyThrAsnPheAsnValAspSerIleTyr 324
1219 GCCGACGCTTTAAAAAATTAGCTATACGAGCAATATTTATTCAGCA 1268
      ||| |||
      ||| |||
325 AlaAspLysValLeuAspGlySerAlaAspValGlnAsnLeuLysVal 340
1269 ATATCAGCGGGGAACAGGACTATTGATCGATCGCTACT..... 1308
      ||| |||
      ||| |||
341 .....AsnGlyValGlyAlaLysGlyAlaSerValAlaIleThrAla 355
1309 ..GCAATTAATACCGCATTTGGCCGCTATGCTGGTGTGCTGCAGCT 1356
      ||| |||
      ||| |||
355 sPlySrlleGlnThrLeuAsnLeuAsnThrThrGlySerGlnSerPheVal 371
1357 GCAGCCGATTTACATTTGAAAGTAACATTAATCTTGCATCACGAA 1406
      ||| |||
      ||| |||
372 SerAlaAspValAlaSerIleSerValLysGlyAsnAlaLysLeuSerLe 388
1407 TAGCAAAAAAGACAAAGTGCACATTCAAAACCTGCTCCGAGAGCGTATT 1456
      ||| |||
      ||| |||
388 uAlaThrGlyAlaIleAspThrThrLeuAsp.....AlaSerSerP 402
1457 TTGCTAAGAAGGCTAATTAATAAGCACTAAAGATGAGAAATGCA 1506
      ||| |||
      ||| |||
402 heGlyGlyAlaLeuAspAlaAspLeuSerThr...SerAlaSerValThr 417
1507 GAATCATCGCTCAAAATGCG...GAGCGGATCACCTCAACGAAGTT... 1551
      ||| |||
      ||| |||
418 SerIleLysGlyLysGlnLysAsnAspLysIleThrIleLysAspValAl 434
1552 .....GATGACTTATGCA 1567
      ||| |||
      ||| |||
434 aValAsnValAlaIleAspGlyGlyAlaGlyAsnAspGluLeuValIleL 451
1568 AAGGTACGGCAAAATTAACCAAGATGAGCTATCAAAAGTTGTGATAC 1617
      ||| |||
      ||| |||
451 ySgIySer.....ThrAlaAspThrleuGlnProThrleuThrAsn 464
1618 TATGAA.....TTGCT 1628
      ||| |||
      ||| |||
465 lIleGluLysValThrIleAspGlyAsnThrLysAspLeuThrleuSerLe 481
```

```

1629 CAACATAGCAAAAAATGTGACAAC...ACCTAGATAGTAATCTCAT 1675
||||| : : : : : ||||| : : : : :
481 uLysylsAlaGlnSerValThrGlnLeuSerPheLysAsnIleAlaLysT 498
||||| : : : : : ||||| : : : : :
1676 CTGTAGCTGATTTACTCTCTCTAATGATCGAGAAATGATTATAGT... 1722
||||| : : : : : ||||| : : : : :
498 hrValThrGlnSerAsnGlyAsnValGlnThrValAlaSerIleLeuAlaAsn 514
||||| : : : : : ||||| : : : : :
1723 .....GCTCACTTCATGTGATGATCAAGATTATCTCTCT 1760
||||| : : : : : ||||| : : : : :
515 AsnAlaThrSplysAlaValThrIleAsnAspGlnSerLeuLysThrI 531
||||| : : : : : ||||| : : : : :
1761 TCAATTTGCTAGGAGGATCTCAGCATTTGAGCTAGCGCTCGCCCTGGCA 1810
||||| : : : : : ||||| : : : : :
531 eAsnPheSerAspValAspAspLys.GlyAlaSerValAlaAlaLysGly 547
||||| : : : : : ||||| : : : : :
1811 GCGGTTCTCAGATTTGAGCTAGCGCTCGCGTCCGGTGGCTCTAGCCAG 1860
||||| : : : : : ||||| : : : : :
548 LysIleValAlaAspLysAlaThrGln.....LeuThr 558
||||| : : : : : ||||| : : : : :
1861 CATTTGAGCTAGCGCTCGCCCTGGCAGCGGATGCCAAGATTGAGCTA 1910
||||| : : : : : ||||| : : : : :
558 rIleAsnSerAsnLysValThrLeuAlaSerAspAlaValAlaGlnAla 575
||||| : : : : : ||||| : : : : :
1911 CGGCTTCGCTCCGGGTGATCTC.....AGCATTTGAGCT 1945
||||| : : : : : ||||| : : : : :
575 LaAsnAlaThrLysIleAspIleAsnAlaAlaLysAspThrValGlyLeu 591
||||| : : : : : ||||| : : : : :
1946 ACGGCTCGCCCTGGCAGCGGTTCTCAAGATT...GAGCTAGAGGCTG 1992
||||| : : : : : ||||| : : : : :
592 Thr.....LeuGlyGlyValAlaLysLeuThrAspLeuThrValAs 605
||||| : : : : : ||||| : : : : :
1993 CGTCCGGGTGCTCTACCGCATTTGAGCTAGCGCTCGCCCTGGCAG 2042
||||| : : : : : ||||| : : : : :
605 nAsnLysGlyAlaPheAlaLeuThrGlyAlaAsnAlaThrAspLeuAsp 622
||||| : : : : : ||||| : : : : :
2043 CGGTAGCCCAAGATT 2056
||||| : : : : :
622 eValLysAsnLeu 626
||||| : : : : :
seq_name: pir2:I41061
seq_documentation_block:
flagellin - Escherichia coli (strain B38)
C:Species: Escherichia coli
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #extl_change 26-Aug-1999
C:Accession: I41061; S47239
R:Zheleznaia, L.A.; Fedorov, O.V.; Brik, A.F.; Matvienko, N.I.
Biochimika 59, 1621-1637, 1994
A:Title: Nucleotide sequence of the Escherichia coli B38 flagellin gene.
A:Reference number: I41061; MUID:95178623
A:Accession: I41061
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-565 <RES>
A:Cross-references: EMBL:236877; NID:g535044; PIDN:CAA8351.1; PID:g535045
C:Genetics:
A:Gene: fljC
C:Function:
A:Description: Involved in bacterial motility and antigenicity
C:Superfamily: flagellin

```

```

alignment_scores:
Quality: 164.00      Length: 546
Ratio: 0.619        Gaps: 26
Percent Similarity: 48.535      Percent Identity: 22.527

```

```
alignment_block:
US-09-306-689-12 x I41061 ..

```

```
Align seg 1/1 to: I41061 from: 1 to: 565

```

```

469 CGCAATATATTTGCAACAGCTCAACACGATTAGGACGATTCAACCCG 518
||||| : : : : : ||||| : : : : :
15 GlnAsnIleLeuLysAsnGlnSerAlaLeu.....SerSer 28
||||| : : : : : ||||| : : : : :
519 TATTGGCTTACAGAGCGTGCATTTGCTTATCCGCTCCCAAAATTGATA 568
||||| : : : : : ||||| : : : : :
28 rIleGlnArgLeuSerSerGlyLeuArgIleAsnSerAlaLysAspAsp 44
||||| : : : : : ||||| : : : : :
569 AATTGTACAGAAAACTAAGCCAGCCACGATT..... 603
||||| : : : : : ||||| : : : : :
45 .....AlaIleGlyIleAlaIleAlaAsnArgPheThr 55
||||| : : : : : ||||| : : : : :
604 GGTTCCTCCGAAGCATTTGTACAAATGCAATTAAGCCAAACTGTATT 653
||||| : : : : : ||||| : : : : :
56 SerAsnIleLysGlyLeuThrGlnAlaIleArgAsnAlaAsnAspGlyI 72
||||| : : : : : ||||| : : : : :
654 ATCTGCATTCATCTATTATTAGCTCAGTATTTGCTGGAATGATTAG 703
||||| : : : : : ||||| : : : : :
72 eSerValAlaGlnThrThrGlnGlyAlaLeu.....SerGlnIleA 86
||||| : : : : : ||||| : : : : :
704 ATGAGGCTTACAGAAATTAACAGCAACCAACATGCTTGTAAAGTCGC 753
||||| : : : : : ||||| : : : : :
86 snAsnAsnLeuGlnArgIleArgGlnLeuThrValGlnAlaSerThrGly 102
||||| : : : : : ||||| : : : : :
754 TTGGAGCTTAACAAATTCATTAATTGAAATTTGCTAATTCAGTAAAC 803
||||| : : : : : ||||| : : : : :
103 ThrAsn...SerAspSerAspLeuAspSerIleGlnAspIleLysSe 118
||||| : : : : : ||||| : : : : :
804 A...CTTGAGCAATTTGGTGGCAATTACT.....CAATT... 837
||||| : : : : : ||||| : : : : :
118 rArgLeuAspGlnIle...AspArgValSerGlyGlnThrGlnPheAsn 134
||||| : : : : : ||||| : : : : :
838 .....GGTCA...AACTACAA..... 852
||||| : : : : : ||||| : : : : :
134 LyValAsnValLeuAlaLysAspGlySerMetLysIleGlnValGlyAla 150
||||| : : : : : ||||| : : : : :
852 ..... 852
||||| : : : : : ||||| : : : : :
151 AsnAspGlyGlnThrIleThrIleAspLeuLysLysIleAspSerAspTh 167
||||| : : : : : ||||| : : : : :
853 .....AATATCAAGGCTTAGGAGCTTAGGACACA 883
||||| : : : : : ||||| : : : : :
167 rLeuGlyLeuAsnGlyPheAsnValAsnGlyLysGlyThrIleAlaSnL 184
||||| : : : : : ||||| : : : : :
884 AACTCAAAATATGCTGAGCTTGATTAAGCTGCGCTGTTATGATGTT 933
||||| : : : : : ||||| : : : : :
184 yAlaAlaThrIleSerAspLeuAlaThrGlyAlaIleValThrAsn 200
||||| : : : : : ||||| : : : : :
934 ATCTCAGGCTA.....TTATCGGCGC 956
||||| : : : : : ||||| : : : : :
201 SerSerAsnIleValValThrThrLysPheAsnAlaLeuAspAlaAl 217
||||| : : : : : ||||| : : : : :
957 AACAGCTGCACCTTGCTGCAGAT..... 981
||||| : : : : : ||||| : : : : :
217 aThrAlaPheSerLysLeuLysAspGlyAspSerValAlaValAlaLaG 234
||||| : : : : : ||||| : : : : :
982 .....AAAAATGCTTCAACA..... 996
||||| : : : : : ||||| : : : : :
234 LnyLysTyThrTyThrAsnAlaSerThrAsnAspPheThrThrGlnAsnThr 250
||||| : : : : : ||||| : : : : :
997 .....GCTAAAAAGTGGG 1010
||||| : : : : : ||||| : : : : :
251 ValAlaThrGlyThrAlaThrThrAspLeuGlyAlaThrLeuLysAlaAl 267
||||| : : : : : ||||| : : : : :
1011 TGGCGGT.....TTGGAATTTGCAACCAACTGTT. 1041
||||| : : : : : ||||| : : : : :
267 aAlaGlyLysSerGlnSerGlyThrTyThrPheAlaAsnGlyLysVala 284
||||| : : : : : ||||| : : : : :
1042 .....GTAATATTACCAAGCCGTTTCTTCTTAC 1071
||||| : : : : : ||||| : : : : :
284 snPheAspValAspAlaSerGlyAsnIleThrIleGlyGlyGlnLysAla 300
||||| : : : : : ||||| : : : : :
1072 ATTTAGCCCAACGTTTGACAGAGTTTATCTTCAACTGGGCTGTGGC 1121
||||| : : : : : ||||| : : : : :

```



```

301 Pheleu.....ValGlyGlyAlaLeuThrThrAsnAspProthrgl 314
1122 TCGTTTAATGCTTCTACTCTTCTCTT..... 1149
314 yserThrProLathrMetSerSerLeuPheYsAlaAlaAspAspLysa 331
1150 .....GCGATTAGCCCATTAGCATTTGCCGGTATTCCGATTAATTAAAT 1194
331 spAlaAlaLeuInsSerTLeaspPheGlyGlyLysTyrGluPheAla 347
1195 CATTGCAAAAAAGTTTAGAGATTATGCCGACGCTTAAAAAATTAGCTA 1244
348 GlyGlyAlaInsSerThrAsnGlyGlyValLysPheLysAspThr...Va 363
1245 TCACGCGAGTAATTTATTAGCAGAAATATACGCGGACACAGGACTATTG 1294
363 lSerSerAspAlaLeuAlaGlnValLysAlaAspSerThrAlaAsnA 380
1295 ATGCATCGGTTACTGCAATTAAATACCGCATG.....GCCGCTATT 1335
380 snValLysTLeThrPheAsnAsnGlyProLeuSerPheThrAlaSerPhe 396
1336 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385
397 GlnAsnGlyValSerGlySerAlaAlaSer..... 406
1386 ACATAATCTTGTTCATCAGCAATAGCAAAAAAGAGAAAGTACCATTCAA 1435
407 ...AsnAlaAlaTyrTLeaspSerGluGluLeuThrThrThrGluS 422
1436 ACTGGTCCGAGAGGCTGATTTTGGTAAAGAGTGGCTTAATATAAGCA 1485
422 eTyr.....AsnThrAsnTyrSerVal 429
1486 ACTAAAGATGAGAAAAATCGAAGAAATCATCGTCAAAATGCGAGCGAT 1535
430 AspLysAspThrGlyAlaValSerValThrGlyGlySerGly..... 443
1536 CACCTCAAAAGCAAGTTGATGATCTTATCGCAAA.....GGTA 1573
444 ThrGlyLysTyrAlaAlaAsnValGlyAlaGlnAlaTyrValGlyAla 460
1574 ACGGCAAAATTTACCAAGATGAGCTATCAAAAGTCTTATTAACATGAA 1623
460 spGlyLysLeuThrThrAsnThrThrSerThrGly..... 471
1624 TTGCTCAACATAGCAAAAAATGTGACAAACAGCTTAGATTAATTAATC 1673
472 .....SerAlaThrLysAspProLeuAsnAlaLeuAspGlnAlaLeuAl 486
1674 ATCTGTAAAGTCATTACCTCGTCTAATGATTCGAGAAATGATTAGTGG 1723
486 aserTLeaspLysPheArgSerSer.....LeuGlyA 497
1724 CTCCAACTTCATGTTGGATCAAAAGTTTATCTTCTT 1761
497 lAlaGlnAsnArgLeuAspSerAlaValThrAsnLeu 509
seq_name: plr2:I37984
seq_documentation_block:
keratin 9, type I, cytoskeletal - human
N:Alternate names: cytokeratin 9; scatter protein 60k chain, placental
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence.revision 02-Jul-1996 #text.change 29-Oct-1999
C:Accession: I37984; S40307; S41161; B35494; I37943
R:Reis, A.; Hennies, H.C.; Landbein, L.; Digweed, M.; Mischke, D.; Drechsler, M.; Schiroc
Nature Genet. 6, 174-179, 1994
A:Title: Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK).
A:Reference number: I37984; MUID:94214498
A:Accession: I37984
A:Status: preliminary; translated from GB/EMBL/DBJ

```

```

A:Molecule type: DNA
A:Residues: 1-622 <RES>
A:Cross-references: EMBL:X75015; NID:9453154; PIDN:CAA52924.1; PID:9453155
R:Landbein, L.; Heid, H.W.; Moill, I.; Franke, W.W.
Differentiation 55, 57-71, 1993
A:Title: Molecular characterization of the body site-specific human epidermal cytoker
A:Reference number: I37943; MUID:94131202
A:Accession: S40307
A:Molecule type: mRNA
A:Residues: 1-622 <LAN>
A:Cross-references: EMBL:Z29074; NID:9435475; PIDN:CAA82315.1; PID:9435476
A:Accession: S77921
A:Molecule type: protein
A:Residues: 29-53; 62-104; 167-188; 199-233; 241-249; 295-362; 374-430; 450-480; 579-604 <LAF
R:Landbein, L.
submitted to the EMBL Data Library, December 1993
A:Reference number: S41161
A:Accession: S41161
A:Molecule type: mRNA
A:Residues: 1-11, 'SR', 13-622 <LAN>
A:Cross-references: EMBL:Z29074; NID:9435475; PIDN:CAA82315.1; PID:9435476
R:Rosen, E.M.; Meromsky, L.; Romero, R.; Selter, E.; Goldberg, I.
Biochem. Biophys. Res. Commun. 168, 1082-1088, 1990
A:Title: Human placenta contains an epithelial scatter protein.
A:Reference number: A35494; MUID:90267446
A:Accession: B35494
A:Molecule type: protein
A:Residues: 'X', 450-465 <ROS>
C:Genetics:
A:Gene: GDB:KRT9; EPPK
A:Cross-references: GDB:303970; OMIM:144200
A:Map position: 17q12-17q21
A:Introns: 213/3; 241/2; 293/3; 347/3; 389/3; 464/2
A:Note: defects in this gene may cause epidermolytic palmoplantar keratoderma
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament
F:1-153/Domain: head #status predicted <HEA>
F:154-459/Domain: helical rod #status predicted <ROD>
F:460-622/Domain: tail #status predicted <TAI>

alignment_block:
US-09-306-689-12 x I37984
Align seg 1/1 to: I37984 from: 1 to: 622

alignment_scores:
Quality: 163.00 Length: 741
Ratio: 0.487 Gaps: 33
Percent Similarity: 45.209 Percent Identity: 20.648

seq_documentation_block:
22 TCTCAGCATTTGAGCTACGCGCTCGCCCTGCGACGCT.....TCTCA 65
|||||: ||| ||| ||| |||: |||
32 serArgPheSerSerSerGlyArgGlyGlyGlyArgPheSerSe 48
|||||: ||| ||| ||| |||: |||
66 ACATTGGAGCTACGCGCTCGCTCGCGGTGCTTACGCCAGCTTGAGCT 115
|||||: ||| ||| ||| |||: |||
48 rSerSerGlyTyrGly.....GlyGlySerSerArgValGlyGlyA 62
|||||: ||| ||| ||| |||: |||
116 ACGCGCTGCGCCCTGCGACGCGTACCAAGATTGGACGACGCGCTGCT 165
|||||: ||| ||| ||| |||: |||
62 rgGly.....GlyGlyGlySerPheGlyTyrSerTyrGlyGly 75
|||||: ||| ||| ||| |||: |||
166 CCGGCTGATCT...CAGCATTTGAGCTACGCGCTCGCGCCCTGCGACGCG 212
|||||: ||| ||| ||| |||: |||
76 serGlyGlyGlyPheSerAlaSerSerLeuGlyGlyGlyPheGlyGly 92
|||||: ||| ||| ||| |||: |||
213 TTCTCAAGTTTGAGCTACGCGCTCGCGGTGCTTACGCCAGCTT 262
|||||: ||| ||| ||| |||: |||
92 yserArgGlyPheGly.....GlyAlaSerGlyGlyGlyTyrSerSerSerg 108
|||||: ||| ||| ||| |||: |||
263 GGAGCTACGCGCTGCGCCCTGCGACGCGTACCAAGATTGGAGCTACGCG 312
|||||: ||| ||| ||| |||: |||

```

108 LysGlyPheGlyGlyGlyPheGlyGlySer.....GlyGlyGly 121  
 313 CTGCTGCTCCGGT...GGATCTAGCTTCCCAAAACCTGGGCAAAAAAAT 359  
 122 PheGlyGlyGlyGlyGlySerGlyPheGlyGlyGlyGly..... 135  
 360 TATCTCTATATTTCCCAAAATACCAATATGATCTGAAACAAGTAATG 409  
 136 .....PheGlyGlyGlyAlaGlyGlyG 143  
 410 GTTACAGATTAGTCAAAAGCCGCAAGATTGGGATTGAGTCA 459  
 143 LysPheGlyGlyLeuThrAlaSerMet.....SerThrMet 156  
 460 AGAAGAAGCAATATATTGCAACAGCTCAAAACGTTTAGGCAGAT 509  
 157 GlnGluLeuSerArgLeuAlaSer..... 165  
 510 TCAACCGCTATTGGCTTAACAGCGCTGCAATGTGTATCCGCTCCAC 559  
 166 .....T 166  
 560 AAATTGATAATGCTACAGAAACCTAACAGCCAGCATTAAGTTCT 609  
 166 LysLeuSplysVal.....GlnAlaLeuGlu 175  
 610 GCGAAAGCAATTGTACAAATATGAATTAAGCCAAACCTGATTATCTGG 659  
 176 AlaSerMet..... 178  
 660 CATTCATCTATTATTAGCTCAGTATTGGCTGGAATGATTAGATGAG 709  
 179 .....AspLeuGlnSer 183  
 710 CCTACAGAAATACAGAACCAACATGCTCTGCTAAGCTGGCTTGAG 759  
 183 LysIleGlnAspTrpLysSplysGlyProAla..... 194  
 760 CTACAAATTCATTAAATGAAATATGCTAATTCAGTAAACACTTGA 809  
 195 .....AlaIleGlnLysSerTrpSerProTyrTrpAsnThrIleAs 208  
 810 CGAATTGTGTCAGCAATATGCAATTTGGTCAAACTACAAATATCA 859  
 208 PaspLeuLysAspGlnIleValAspLeu..... 217  
 860 AAGCTTAGGACTTTAGAGACAACTCAAAATATCGGTGACCTGAT 909  
 218 .....ThrValGlyAsnSer...LysThrLeuLeuAspIleAsp 229  
 910 AAGCTGCTGCTTGTAGATTTATCTCAGGCTATTATCGGGCGCAC 959  
 230 AsnThrArgMetThrLeuAspAspPheArg..... 239  
 960 AGCTGCACTTGTACTTCAGATTAATGCTTCAACAGCTAAATAAGTGG 1009  
 240 .....IleLysPheGlnMetGlnAsnLeuArgGlnGlyAla 253  
 1010 GTGCGGCTTTGAATTGGCAAAACCACTTTGGTAATATTACCAAGCC 1059  
 253 sPAlaAspIleAsnGlyLeuArgGlnValLeuAspAsnLeuThr..... 267  
 1060 GTTCTCTTACATTTTAGCCCAACGTTGCAAGCTTATCTTCAAC 1109  
 267 ..... 267  
 1110 TGGGCTGTGCTGCTTAAATGCTTCTACTGTTCTTGCATTAAGCC 1159  
 268 .....MetGlnLysSerAspLeuGlnMetGlnTyrGlu 279  
 1160-CATTAGCATTTGCCGGTATTGCCGATTAATTAATCAAGCAAAAGTTTA 1209  
 279 hrLeuGlnGlnGluLeuMetAlaLeuLysLysAsnHisLysGlnGluMet 295

1210 GAGAGTATCCGACAGCTTTAAAAAATATGCTATGACGAGATATTT 1259  
 296 SerGlnLeuThr.....GlyGlnSer... 302  
 1260 ATTAGCAGAAATATACGGGGAAACAGGACTATTGATGATCGTTACTG 1309  
 303 .....SerGlyAspValAsnValGlnIleAsn 312  
 1310 CAATTATACCGCATTTGCCGCTATTGCTGTGTCTGCTGCTGCA 1359  
 312 al.....AlaProGly 315  
 1360 GCGCATTTAATATTGAAAAAGTTAAACATATCTGTCTACAGATAG 1409  
 316 LysAspLeuThr...LysThrLeuAsnAspMet.....ArgGln 327  
 1410 CAAAAAGAGAAAGTGACCATTTCAAAACCTGTTCCGAGAGCTGATTG 1459  
 327 ngLutyrGlnGlnIleuIleAlaLysAsn.....A 337  
 1460 CTAAAGAGTGCCTAATTATTAAGCAACTAAAGATGAGAAATCGAAGA 1509  
 337 rglLysAspIleGlnAsnGlnTyrGlnThrGlnIleThrGlnIleGlnHis 353  
 1510 ATCATCGCTCAAAATGCGCGAGCATCACCTCA.....AAGCAAGTTGA 1553  
 354 GluValSerSerSerGlyGlnGluValGlnSerSerAlaLysGlnVal 370  
 1554 TGATCTTATTCGCAAAAGTAAAGCGCAAAAT.....ACCCAAGATG 1594  
 370 rglLeu...ArgHisGlyValGlnGlnLeuGlnIleGlnGlnSerG 386  
 1595 AGCTATCAAAAGTTGTGATAACTATGAAATGCTCAACATAGCAAAAT 1644  
 386 LysLeuSerLysAlaAlaLeuGlnLysSerLeuGlnAspTrpLysAsn 402  
 1645 .....GTGACAAACAGCTTAGATTAATGATTCATCTCTAAG 1682  
 403 ArgTyrCysGlyGlnLeuGlnMetIleGlnGlnIleSerMetLeuGln 419  
 1683 TGCA.....TTTACCTGCTTAATGATTCGA 1708  
 419 uAlaGlnIleThrAspValArgGlnGlnGlnGlnGlnGlnGlnGln 436  
 1709 GAATGATATTAGTGCCTCCAACTCAATGTTGATGATCAAAATATCTCT 1758  
 436 LysSerLeuLeuSerIleLysMetArgLeuGlnLysGlnIleGluThr 452  
 1759 CTTCAA..... 1764  
 453 TyrHisAsnLeuLeuGlnGlnGlnGlnLysAspPheGlnSerSerGlyAla 469  
 1765 .....TTTGTAGAGGATCTCAGATTTGAGAGCTACGCGC 1798  
 469 agLysIleGlyLeuGlnGlyArgGlySerGlySerTyrGly 485  
 1799 TGGGCTGCGCAGCGTTCTCAAGATTGGAGCTACGCGCTCCGGGT 1848  
 486 .ArgLysSerArgGlySerGlySerGlySerTyrGlySerGlySer 501  
 1849 GCGCTTAGCCAGCATTTGAGCTACGCGCTGCGCTGCGAGCGTACCA 1898  
 502 GlyGlyGlyTyrGlyGlyGlySerGlySerArgGlySerGlyGly 517  
 1899 AGATTGAGCTACGCGCTGCGCTGCGGTGAGATCAAGATTTGAGATGAG 1948  
 518 .SerTyrGlyGlyGlySerGlySerGlySerGlySerGlyGlyGly 534  
 1949 GCGTGGCGCTGCGAGCGGTTCTCAAGATTGAGCTACGCGCTGCGTCCG 1998  
 534 LysLysLysSerGlyGlyGlnHisSerGlyGlySerGlyGlyGlnHisSer 550



[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:36 ; Search time 47.48 Seconds  
(Without alignments)  
18,532 Million cell updates/sec

Title: US-09-306-689-11

Perfect score: 290  
Sequence: 1 QHMSYGLRGSGSGQDMSYGL.....SYGLRPGSGSGQDMSYGLRPG 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Issued Patents AA:\*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	100.0	49	1 US-08-387-156-4	Sequence 4, Appli
2	290	100.0	49	2 US-08-694-865-4	Sequence 4, Appli
3	290	100.0	49	3 US-08-878-748-4	Sequence 4, Appli
4	290	100.0	49	4 US-09-124-491-4	Sequence 4, Appli
5	290	100.0	544	1 US-08-387-156-10	Sequence 10, Appli
6	290	100.0	544	2 US-08-694-865-10	Sequence 10, Appli
7	290	100.0	544	3 US-08-878-748-10	Sequence 10, Appli
8	290	100.0	544	3 US-09-124-491-10	Sequence 10, Appli
9	290	100.0	699	3 US-08-694-865-16	Sequence 16, Appli
10	290	100.0	699	3 US-09-124-491-16	Sequence 16, Appli
11	290	100.0	977	1 US-08-387-156-8	Sequence 8, Appli
12	290	100.0	977	2 US-08-694-865-8	Sequence 8, Appli
13	290	100.0	977	2 US-08-878-748-8	Sequence 8, Appli
14	290	100.0	977	3 US-09-124-491-8	Sequence 8, Appli
15	187.5	64.7	44	1 US-07-690-983D-45	Sequence 45, Appli
16	187.5	64.7	84	1 US-07-690-983D-47	Sequence 47, Appli
17	152	52.4	55	3 US-08-458-814-7	Sequence 7, Appli
18	150.5	51.9	52	3 US-08-458-814-6	Sequence 6, Appli
19	104.5	36.0	24	1 US-07-690-983D-43	Sequence 43, Appli
20	96	33.1	20	1 US-07-690-983D-40	Sequence 40, Appli
21	74	25.5	90	1 US-08-341-219-19	Sequence 19, Appli
22	72	24.8	18	1 US-07-690-983D-20	Sequence 20, Appli
23	70	24.1	16	1 US-08-453-588-26	Sequence 26, Appli
24	70	24.1	16	3 US-08-521-079-26	Sequence 26, Appli
25	69	23.8	17	1 US-07-690-983D-16	Sequence 16, Appli
26	67	23.1	14	1 US-08-453-588-23	Sequence 23, Appli
27	67	23.1	14	3 US-08-521-079-23	Sequence 23, Appli
28	67	23.1	936	1 US-08-455-970A-12	Sequence 12, Appli

29	66.5	22.9	89	1	US-08-341-219-20	Sequence 20, Appli
30	66	22.8	14	1	US-07-690-983D-24	Sequence 24, Appli
31	66	22.8	18	1	US-07-690-983D-28	Sequence 28, Appli
32	66	22.8	26	1	US-08-446-692-29	Sequence 29, Appli
33	66	22.8	26	2	US-08-488-351A-29	Sequence 29, Appli
34	66	22.8	750	5	5457037-3	Patent No. 5457037
35	66	22.8	751	5	5457037-5	Patent No. 5457037
36	65.5	22.6	17	1	US-07-690-983D-18	Sequence 18, Appli
37	65	22.4	17	1	US-08-188-223-5	Sequence 5, Appli
38	65	22.4	17	3	US-08-968-466-5	Sequence 5, Appli
39	65	22.4	29	1	US-08-446-692-28	Sequence 28, Appli
40	65	22.4	29	1	US-08-488-351A-28	Sequence 28, Appli
41	65	22.4	34	1	US-08-446-692-16	Sequence 16, Appli
42	65	22.4	34	2	US-08-488-351A-16	Sequence 16, Appli
43	65	22.4	37	2	US-08-446-692-24	Sequence 24, Appli
44	65	22.4	37	2	US-08-488-351A-24	Sequence 24, Appli
45	64	22.1	12	1	US-08-453-588-24	Sequence 24, Appli

# ALIGNMENTS

RESULT 1  
US-08-387-156-4  
Sequence 4, Application US/08387156  
Patent No. 5723129  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: GPRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/387,156  
FILING DATE: 10-FEB-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-387-156-4  
Query Match 100.0%; Score 290; DB 1; Length 49;  
Best Local Similarity 100.0%; Pred. No. 5.3e-27;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 OHMSYGLRPGSGSQDMSYGLRPGSSQHMYSYGLRPGSGSQDMSYGLRPG 49  
DB 1 OHMSYGLRPGSGSQDMSYGLRPGSSQHMYSYGLRPGSGSQDMSYGLRPG 49

## RESULT 2

US-08-694-865-4  
Sequence 4, Application US/08694865

Patent No. 5837268

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: MANNIS, JOHN G.

TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: REED &amp; ROBINS LLP

STREET: 285 HAMILTON AVENUE, SUITE 200

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/694,865

FILING DATE: 09-AUG-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: MCCracken, THOMAS P.

REGISTRATION NUMBER: 38,548

REFERENCE/DOCKET NUMBER: 9001-0016.22

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415)327-3400

TELEFAX: (415)327-3231

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 49 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-694-865-4

Query Match 100.0%; Score 290; DB 2; Length 49;  
Best Local Similarity 100.0%; Pred. No. 5.3e-27;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OHMSYGLRPGSGSQDMSYGLRPGSSQHMYSYGLRPGSGSQDMSYGLRPG 49  
DB 1 OHMSYGLRPGSGSQDMSYGLRPGSSQHMYSYGLRPGSGSQDMSYGLRPG 49

## RESULT 3

US-08-878-748-4  
Sequence 4, Application US/08878748

Patent No. 5969126

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: REDMOND, MARK J.

APPLICANT: HUGHES, HUW P.A.

TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: REED &amp; ROBINS

STREET: 635 BRYANT STREET

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 94301

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/878,748

FILING DATE: 19-JUN-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/387,156

FILING DATE: 10-FEB-1995

APPLICATION NUMBER: US 07/960,932

FILING DATE: 14-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/779,171

FILING DATE: 16-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: ROBINS, ROBERTA L.

REGISTRATION NUMBER: 33,208

REFERENCE/DOCKET NUMBER: 9001-0016.21

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 617-8999

TELEFAX: (415) 327-3231

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 49 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-878-748-4

Query Match 100.0%; Score 290; DB 2; Length 49;  
Best Local Similarity 100.0%; Pred. No. 5.3e-27;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OHMSYGLRPGSGSQDMSYGLRPGSSQHMYSYGLRPGSGSQDMSYGLRPG 49  
DB 1 OHMSYGLRPGSGSQDMSYGLRPGSSQHMYSYGLRPGSGSQDMSYGLRPG 49

## RESULT 4

US-09-124-491-4  
Sequence 4, Application US/09124491

Patent No. 6022960

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: MANNIS, JOHN G.

TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: REED &amp; ROBINS LLP

STREET: 285 HAMILTON AVENUE, SUITE 200

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/124,491

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/694,865

FILING DATE: 09-AUG-1996

APPLICATION NUMBER: US 08/387,156

FILING DATE: 10-FEB-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-4

Query Match 100.0%; Score 290; DB 3; Length 49;  
Best Local Similarity 100.0%; Pred. No. 5.3e-27;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSSQDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 49  
DB 1 OHMSYGLRPGSSQDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 49

RESULT 5  
US-08-387-156-10  
Sequence 10, Application US/08387156  
Patent No. 5723129  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/387,156  
FILING DATE: 10-FEB-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-387-156-10

Query Match 100.0%; Score 290; DB 1; Length 544;  
Best Local Similarity 100.0%; Pred. No. 8.1e-26;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSSQDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 49  
DB 494 OHMSYGLRPGSSQDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 542

RESULT 6  
US-08-694-865-10  
Sequence 10, Application US/08694865  
Patent No. 5837268  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200.  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,865  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-865-10

Query Match 100.0%; Score 290; DB 2; Length 544;  
Best Local Similarity 100.0%; Pred. No. 8.1e-26;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSSQDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 49  
DB 494 OHMSYGLRPGSSQDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 542

RESULT 7  
US-08-878-748-10  
Sequence 10, Application US/08878748  
Patent No. 5969126  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/878,748  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016, 21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-878-748-10

Query Match 100.0%; Score 290; DB 2; Length 544;  
Best Local Similarity 100.0%; Pred. No. 8,1e-26;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
DB 494 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 542

RESULT 8  
US-09-124-491-10  
Sequence 10, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016, 22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-10

Query Match 100.0%; Score 290; DB 3; Length 544;  
Best Local Similarity 100.0%; Pred. No. 8,1e-26;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
DB 494 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 542

RESULT 9  
US-08-694-865-16  
Sequence 16, Application US/08694865  
Patent No. 5837268  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,865  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016, 22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 699 amino acids



TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-865-16

Query Match 100.0%; Score 290; DB 2; Length 699;  
Best Local Similarity 100.0%; Pred. No. 1,1e-25;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQMSYGLRPGSSQDMSYGLRPG 49  
|||||  
Db 9 OHMSYGLRPGSGSDMSYGLRPGSSQMSYGLRPGSSQDMSYGLRPG 57

RESULT 10  
US-09-124-491-16  
Sequence 16, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991

ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:  
LENGTH: 699 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-16

Query Match 100.0%; Score 290; DB 3; Length 699;  
Best Local Similarity 100.0%; Pred. No. 1,1e-25;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQMSYGLRPGSSQDMSYGLRPG 49  
|||||

Db 9 OHMSYGLRPGSGSDMSYGLRPGSSQMSYGLRPGSSQDMSYGLRPG 57

RESULT 11  
US-08-387-156-8  
Sequence 8, Application US/08387156  
Patent No. 5723129  
GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P. A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/387,156  
FILING DATE: 10-FEB-1995  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-387-156-8

Query Match 100.0%; Score 290; DB 1; Length 977;  
Best Local Similarity 100.0%; Pred. No. 1,1e-25;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQMSYGLRPGSSQDMSYGLRPG 49  
|||||  
Db 927 OHMSYGLRPGSGSDMSYGLRPGSSQMSYGLRPGSSQDMSYGLRPG 975

RESULT 12  
US-08-694-865-8  
Sequence 8, Application US/08694865  
Patent No. 5837268  
GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO

STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,865  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-865-8

Query Match 100.0%; Score 290; DB 2; Length 977;  
Best Local Similarity 100.0%; Pred. No. 1.6e-25;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 49  
DB 927 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 975

RESULT 13  
US-08-878-748-8  
Sequence 8, Application US/08878748  
Patent No. 5969126  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: REDMOND, MARK J.  
APPLICANT: HUGHES, HOW P.A.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/878,748  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.

REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 9001-0016.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-878-748-8

Query Match 100.0%; Score 290; DB 2; Length 977;  
Best Local Similarity 100.0%; Pred. No. 1.6e-25;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 49  
DB 927 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 975

RESULT 14  
US-09-124-491-8  
Sequence 8, Application US/09124491  
Patent No. 6022960  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/124,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/387,156  
FILING DATE: 10-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/960,932  
FILING DATE: 14-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/779,171  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-124-491-8

Query Match	100.0%;	Score 290;	DB 3;	Length 977;
Best Local Similarity	100.0%;	Pred. No. 1.6e-25;		
Matches 49;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 QHWSGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 49  
 |||||  
 Db 927 QHWSGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 975

```

RESULT 15
US-07-690-983D-45
Sequence 45: Application US/07650983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690,983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
Type: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-690-983D-45

```

Query Match	64.7%	Score 187.5	DB 1:	Length 44;
Best Local Similarity	69.4%	Pred. NO. 2.2e-15;		
Matches 34;	Conservative 4;	Mismatches 2;	Indels 9;	Gaps 3;
OY	1	QHWSTGLRPGSGSDPDWSTGLRPGSSQHWSTGLRPGSGSDPDWSTGLRPG	49	
db	3	EHWSYGLRPG---EHWSYGLRPG---EHWSYGLRPG	42	

Search completed: March 2, 2001, 10:54:37  
Job time: 602 sec

**THIS PAGE BLANK (USPTO)**



OY 46 LRP 48  
DB 87 LRP 89

## RESULT 2

O9VTR6 PRELIMINARY: PRT: 836 AA.  
AC O9VTR6;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE ZINC FINGER PROTEIN RIN ZF.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99175194; PubMed=10075714;  
RA Tillotson L.G.;  
RT "RIN ZF, a novel zinc finger gene, encodes proteins that bind to the  
RT CACC element of the gastrin promoter";  
RL J. Biol. Chem. 274:8123-8128(1999).  
DR EMBL; AF091457; AAD2522.1; -.  
DR HSSP; P07248; IARD.  
DR INTERPRO: IPR000210; -.  
DR INTERPRO: IPR000822; -.  
DR PFAM; PF00096; zf-C2H2; 2.  
DR PFAM; PF00651; BTF; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; UNKNOWN\_1.  
SQ SEQUENCE 836 AA; 91698 MW; E4F48AB2720676C1 CRC64;

Query Match 32.6%; Score 94.5; DB 11; Length 836;  
Best Local Similarity 47.8%; Pred. No. 0.0096;  
Matches 22; Conservative 7; Mismatches 14; Indels 3; Gaps 3;

OY 3 WSGLRPGSSODMSYGLRPGSSODMSYGLRPG 48  
DB 626 FKYGMPGT-SSDFKYGILP-STSNFKYGILPGA-PNDFKYGILP 668

## RESULT 3

O9VTR6 PRELIMINARY: PRT: 1729 AA.  
AC O9VTR6;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE PERICARDINE.  
GN PRC.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OREGON R;  
RA Chatterjee A., Astler M., Zaffran S., Semeriva M., Gratecos D.;  
RT "Characterization of an extracellular matrix component specifically  
RT expressed in the basal lamina of the dorsal vessel in Drosophila  
RT melanogaster";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases..  
DR EMBL; AF203342; AAF13280.1; -.  
SQ SEQUENCE 1729 AA; 164661 MW; DA9B1B7FABSCDEDB CRC64;

Query Match 27.9%; Score 81; DB 5; Length 1729;  
Best Local Similarity 37.9%; Pred. No. 0.59; Indels 14; Gaps 4;  
Matches 22; Conservative 8; Mismatches 14; Indels 14; Gaps 4;

OY 5 YGLRPGSSODMS-----YGLRPG-----GSSQHWSYGLRPGSSODMS-----YGLRPG 49  
DB 400 YGSGPGIGGTGAGCPGYGTGPGIGGTGAGCP-GVGTGPGTATGAGCPGYGTGPG 456

## RESULT 4

O9VTR6 PRELIMINARY: PRT: 1039 AA.  
AC O9VTR6;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE CG5700 PROTEIN.  
GN CG5700.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Burtis K.C., Burtis D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gehlert W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyev C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003543; AAF49980.1; -.  
DR FLYBASE; FBgn0036231; CG5700.  
SQ SEQUENCE 1039 AA; 102635 MW; 6A710871139BD01F CRC64;

Query Match 27.8%; Score 80.5; DB 5; Length 1039;  
Best Local Similarity 40.7%; Pred. No. 0.4; Indels 11; Gaps 4;  
Matches 22; Conservative 6; Mismatches 15; Indels 11; Gaps 4;

OY 5 YGLRPGSSODMSYGLRPG-----GSSQHWSYGLRPGSSODMS-----YGLRPG 49



AC 097643;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE FIBRINOGEN A ALPHA CHAIN (FRAGMENT).  
 OS Lama glama (Llama).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
 OX NCBI\_TaxID=9844;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Murakawa M.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF095464; AAC67563.1; -  
 DR HSSP; P02671; 1FZG.  
 FT NON\_TER 1  
 FT SEQUENCE 462 AA; 48968 MW; 377B503CF32E698 CRC64;  
 SQ

Query Match  
 Best Local Similarity 25.3%; Score 73.5; DB 6; Length 462;  
 Matches 19; Conservative 9; Mismatches 17; Indels 5; Gaps 5;

QY 3 WSYGL-RPGSGSDMSYGLRPGSSSQHWSYG-LRPGSGSDMSYGL-RPG 49  
 Db 166 WTHGSEPPGSAG-TWTHGSEPPGSAGWTHGLEPPGSAG-SWTHGSEPPG 213

RESULT 9  
 ID 013344 PRELIMINARY; PRT; 528 AA.  
 AC 013344;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE FUS-LIKE PROTEIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-CARCINOMA;  
 RA Itch K., Kawase M.;  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROBABLE RNA-BINDING PROTEIN (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNP).  
 DR EMBL; U36561; AAA79948.1; -  
 DR INTERPRO: IPR000504; -  
 DR INTERPRO: IPR001419; -  
 DR INTERPRO: IPR001876; -  
 DR INTERPRO: IPR002961; -  
 DR PFM; PFO0076; rtm; 1.  
 DR PFM; PFO0641; zf-RanBP; 1.  
 DR PRINTS: PRO00210; GLUTENIN.  
 DR PRINTS: PRO1237; TNFC.  
 DR PROSITE: PS00030; RNP\_1, FALSE\_NEG.  
 KW RNA-binding; Nuclear protein.  
 FT DOMAIN 1 180  
 FT AC 159 164  
 FT FT 163 168  
 FT DOMAIN 179 199  
 FT FT 181 199  
 FT DOMAIN 205 212  
 FT FT 205 212  
 FT DOMAIN 226 234  
 FT FT 226 234  
 FT DOMAIN 289 294  
 FT FT 289 294  
 FT DOMAIN 337 344  
 FT FT 337 344  
 FT DOMAIN 374 527  
 FT FT 374 527  
 FT DOMAIN 396 412  
 FT FT 396 412  
 FT DOMAIN 416 422  
 FT FT 416 422

FT ZN\_FING 431 450 C4-TYPE (POTENTIAL).  
 FT NON\_TER 528  
 SQ SEQUENCE 528 AA; 53377 MW; 43CFD6CBDA57A258 CRC64;

Query Match  
 Best Local Similarity 24.8%; Score 72; DB 4; Length 528;  
 Matches 20; Conservative 2; Mismatches 12; Indels 16; Gaps 2;

QY 5 YGLRPGSGSDMSYGLRPGSSSQHWSYG-LRPGSGSDMSYGLRPG 45  
 Db 99 YGPQPTPSSYSGSYSSSOTSSTGQPP-----STGQQPSYSGGQDSYG 141

RESULT 10  
 ID 091020 PRELIMINARY; PRT; 701 AA.  
 AC 091020;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HYPOTHETICAL 73.3 KDA PROTEIN.  
 GN SCL6.26C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmid and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL MOL. Microbiol. 21:77-96(1996).  
 DR EMBL; AL159139; CAB76891.1; -  
 KW Hypothetical protein  
 SQ SEQUENCE 701 AA; 73279 MW; 55BDFCBCT77BC412 CRC64;

Query Match  
 Best Local Similarity 24.8%; Score 72; DB 2; Length 701;  
 Matches 21; Conservative 4; Mismatches 19; Indels 10; Gaps 3;

QY 5 YGLRPGSGSDMSYGLRPGSSSQHWSYG-LRPGSGSDMSYGLRPG 49  
 Db 577 HGARDGHRGDDRGARPGHGTGDDGGDHRSDGARGDYGARGP-GYCAODG 629

RESULT 11  
 ID 057480 PRELIMINARY; PRT; 1068 AA.  
 AC 057480;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE PHOTORECEPTOR GUANYLATE CYCLASE 1 (FRAGMENT).  
 GN GCL.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;



```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RHODE ISLAND RED; TISSUE-RETINA-PIGMENT EPITHELIUM-CHOROIO;
RA Semple-Rowland S.L., Lee N.R., Van Hooser J.P., Palczewski K.,
BA Behr W.;
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(1997).
DR EMBL: AF036942; AAC24500.1; -.
DR HSSP: 002846; IAML.
DR INTERPRO: IPR000719; -.
DR INTERPRO: IPR001054; -.
DR INTERPRO: IPR001245; -.
DR INTERPRO: IPR001828; -.
DR PFAM: PF00069; Kinase; 1.
DR PFAM: PF00211; guanylate cyc; 1.
DR PFAM: PF01094; ANF receptor; 1.
DR PRINTS: PR00109; TRKINASE.
DR PROSITE: PS00452; GUANYLATE_CYCLASES; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KM Lyase.
FT NON_TER
SQ SEQUENCE 1068 AA; 11828 MW; 3FAC693E1597174B CRC64;

Query Match 24.8%; Score 72; DB 13; Length 1068;
Best Local Similarity 39.0%; Pred. No. 3.4;
Matches 16; Conservative 3; Mismatches 14; Indels 8; Gaps 2;

OY 10 GSGSDMS-YGLRPG-----GSSQHSYGLRPGSGSQDW 42
DB 356 GKGDRLMPVGLPEGRGLAYRGHSHVHPHSSPCTDSCGW 396

RESULT 12
OY 09T513 PRELIMINARY; PRT; 107 AA.
AC 09T513;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CHICKEN LUTEINIZING HORMONE-RELEASING HORMONE II (FRAGMENT).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Adler L.A., Sherwood N.M., Grendell R.L., Golos T.G., Terasawa E.;
RT "cDNA of a second form of luteinizing hormone releasing, chicken LHRH-
RT II, isolated from the non-human primate brain (Abstract 632.8).";
RL Abstr. Soc. Neurosci. 24:1607-1607(1998).
DR EMBL: AF104307; AAD13775.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
FT NON_TER
SQ SEQUENCE 107 AA; 11823 MW; FACE52703C3CB1D CRC64;

Query Match 24.5%; Score 71; DB 6; Length 107;
Best Local Similarity 46.2%; Pred. No. 0.44;
Matches 16; Conservative 3; Mismatches 12; Indels 6; Gaps 2;

OY 1 QHWSYGLRPG-----SGSQDMSYGLR-PGSSQHSYGL 33
DB 25 QHWSHGVPKGRKALSSADPNALRPPAGSPAQATYGL 63

RESULT 13
OY 097655 PRELIMINARY; PRT; 114 AA.
AC 097655;
DT 01-MAY-1999 (Tremblrel. 10, Created)

```

```

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
DE PRECURSOR.
GN GNRH2.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA White R.B., Urbanski H.F., Fernald R.D.;
RT "A second gene for gonadotropin-releasing hormone is expressed in the
RT Rhesus macaque (Abstract #632.18).";
RL Abstr. Soc. Neurosci. 24:1609-1609(1998).
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC POLICIE-STIMULATING HORMONES.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
DR EMBL: AF097356; AAD09106.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
DR PRODOM: PD029787; -. 1.
KM Hormone; Amidation.
SQ SEQUENCE 114 AA; 12533 MW; 8B70D690D5BD5103 CRC64;

Query Match 24.5%; Score 71; DB 6; Length 114;
Best Local Similarity 46.2%; Pred. No. 0.47;
Matches 18; Conservative 3; Mismatches 12; Indels 6; Gaps 2;

OY 1 QHWSYGLRPG-----SGSQDMSYGLR-PGSSQHSYGL 33
DB 25 QHWSHGVPKGRKALSSADPNALRPPAGSPAQATYGL 63

RESULT 14
OY 09X111 PRELIMINARY; PRT; 257 AA.
AC 09X111;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE F19G14.22 PROTEIN.
GN F19G14.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Strain-CV, COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Rongning C.M., Benito M.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F19G14 genomic sequence.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006438; AAD41986.1; -.
DR HSSP: P00303; 2CBP.
SQ SEQUENCE 257 AA; 27400 MW; DA59B77A93D53A17 CRC64;

```

```

Query Match 24.5%; Score 71; DB 10; Length 257;
Best Local Similarity 26.1%; Pred. No. 1.1;
Matches 18; Conservative 6; Mismatches 19; Indels 26; Gaps 3;

OY 3 WSYGLRPG-----SGSQDMSYGLR-PGSSQHSYGLRPGS----- 37
DB 87 WGMGVVPMNTHSSSGSGGCMGCPNNNYSGSGSGSGSGMVGCGHSHKYNVNTATYNGPKTIIV 146

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using 'sw model

Run on: March 2, 2001, 11:01:20 ; Search time 32.03 Seconds  
(Without alignments)  
49,404 Million cell updates/sec

Title: US-09-306-689-11

Perfect score: 290  
Sequence: 1 QHWSYGLRPGSGSDMSYGL.....SYGLRPGSGSDMSYGLRPG 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SWISSPROT\_39.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77	27.2	407	1 SM41_HEMPU	Q26264 hemilectrot
2	79	26.6	445	1 SM50_SPRPU	P11994 strongyloce
3	76.5	26.4	67	1 GONI_MACMU	P55247 macaca mula
4	76.5	26.4	92	1 GONI_HUMAN	P01148 homo sapien
5	74	25.5	90	1 GONI_MOOSE	P13562 mus musculu
6	74	25.5	91	1 GONI_PIG	P49921 sus scrofa
7	74	25.5	92	1 GONI_RAT	P07490 rattus norv
8	71	24.5	414	1 SVS2_RAT	P22006 rattus norv
9	70	24.1	518	1 FUS_MOUSE	P56959 mus musculu
10	69	23.8	512	1 FUS_BOVIN	Q28009 bos taurus
11	69	23.8	526	1 FUS_HUMAN	P35637 homo sapien
12	68.5	23.6	94	1 GONI_HABPU	P51918 haplochromi
13	68	23.4	335	1 SM34_LYTP1	Q05994 lytechinus
14	66.5	22.9	89	1 GONI_XENLA	P45656 xenopus lae
15	66	22.8	776	1 ISOA_PSEAY	P10342 pseudomonas
16	66	22.8	776	1 ISOA_PSEAY	P25501 pseudomonas
17	65	22.4	92	1 GONI_TURGB	Q95335 cupata glis
18	65	22.4	120	1 GONI_HUMAN	Q43535 homo sapien
19	64	22.1	92	1 GONI_CAVPO	O54712 cavia porce
20	63.5	21.9	95	1 GONI_PAGMA	P70074 pagrus majo
21	63.5	21.9	95	1 GONI_SPAU	P51919 sparus aur
22	63	21.7	61	1 GONI_SHEEP	Q28588 ovis aries
23	63	21.7	63	1 GONI_MESAU	Q09153 mesocricetu
24	62	21.4	839	1 GLUT_WHEAT	P10368 triticum ae
25	61	21.0	767	1 ACES_CHICK	P36196 gallus gall
26	60.5	20.9	627	1 K2CI_MOUSE	P41044 mus musculu
27	60	20.7	92	1 GONI_CHICK	P37042 gallus gall
28	60	20.7	379	1 ROA3_HUMAN	P51991 homo sapien
29	60	20.7	505	1 SCRY_SALTY	P22340 salmoneilla
30	59.5	20.5	838	1 GLUT_WHEAT	P08489 triticum ae
31	59	20.3	10	1 GONI_ALUMI	P37041 alligator m
32	59	20.3	114	1 GONI_TURGB	Q95336 cupata glis
33	59	20.3	567	1 ENV_AVISN	P31796 avian splee

34	58	20.0	80	1 GONI_CLAGA	P33439 ciarias gar
35	58	20.0	82	1 GONI_SALTR	P45653 salmo trutt
36	58	20.0	370	1 ALLS_DIPPU	P12764 diptoptera
37	57	19.7	385	1 HEM6_SOTBN	P35055 glycine max
38	57	19.7	391	1 HEM6_HORVU	Q42840 hordeum vul
39	57	19.7	492	1 SYNM_YEAST	P25345 saccharomyc
40	57	19.7	1321	1 IRS2_MOUSE	P81122 mus musculu
41	56	19.3	219	1 TRPF_CAUCR	P12289 caulobacter
42	56	19.3	313	1 SYPH_HUMAN	P08247 homo sapien
43	56	19.3	373	1 RO31_XENLA	P51968 xenopus lae
44	56	19.3	385	1 RO32_XENLA	P51992 xenopus lae
45	55.5	19.1	86	1 GONI2_CLAGA	P43306 ciarias gar

## ALIGNMENTS

```
RESULT 1
ID SM41_HEMPU STANDARD; PRT; 407 AA.
AC Q26264;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 41 KDA SPICULE MATRIX PROTEIN PRECURSOR (HSM41) (HPSMC).
OS Hemilectrotus pulcherrimus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinozoa; Echinodermata; Echinodermata;
OC Hemilectrotus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93075615; PubMed=1445780;
RA Katoh-Fukui Y., Noce T., Ueda T., Fujiwara Y., Hashimoto N.,
RA Tanaka S., Higashinakagawa T.;
RT "Isolation and characterization of cDNA encoding a spicule matrix
RT protein in Hemilectrotus pulcherrimus micromeres.";
RL Int. J. Dev. Biol. 36:353-361(1992).
CC -!- FUNCTION: MAJOR MATRIX PROTEIN OF THE SEA URCHIN EMBRYO SPICULE
CC WHICH DIRECTS CRYSTAL GROWTH IN CERTAIN ORIENTATIONS AND INHIBIT
CC GROWTH IN OTHERS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE
CC MICROMERE/PRIMARY MESENCHYME CELLS (PMC) LINEAGE.
CC -!- DEVELOPMENTAL STAGE: APPEARS IN GASTRULAE AND REMAINS AT A SIMILAR
CC LEVEL UNTIL THE PLUTEUS STAGE.
CC -!- DOMAIN: THE REPETITIVE DOMAIN MAY PROVIDE A CALCITE BINDING
CC MATRIX.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE SM50 FAMILY.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collapration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC EMBL: S48755; AAB24285.1; -.
CC INTERPRO: IPR001304; -.
CC PROSITE: PS00615; C-TYPE LECTIN_1; FALSE_NEG.
CC PROSITE: PS50041; C-TYPE LECTIN_2; 1.
CC DR Matrix protein; Signal; Repeat.
CC FT SIGNAL 1 17 POTENTIAL.
CC FT CHAIN 18 407 41 KDA SPICULE MATRIX PROTEIN.
CC FT DOMAIN ? ? C-TYPE LECTIN.
CC FT DOMAIN 104 179 PRO-RICH.
CC SO SEQUENCE 407 AA; 42383 MM; 5A5397C46B80F6B5 CRC64;
Query Match 27.2%; Score 79; DB 1; Length 407;
Best Local Similarity 30.6%; Pred. No. 0.15;
Matches 15; Conservative 11; Mismatches 23; Indels 0; Gaps 0;
```



Db 2 GCSSQHSYGLRPG-GRKD 19

RESULT 4

ID	GONI_HUMAN	STANDARD	PRT	92 AA
AC	P01148			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I) (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED PEPTIDE I].			
DE	GNRH I OR GNRH OR LHRH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-89366682; Pubmed-2671939;			
RA	Hayflick J.S., Adelman J.P., Seeburg P.H.;			
RT	"The complete nucleotide sequence of the human gonadotropin-releasing hormone gene."			
RL	Nucleic Acids Res. 17:6403-6403(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-86094338; Pubmed-2867548;			
RA	Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;			
RT	"Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat."			
RL	Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-85012739; Pubmed-6090951;			
RA	Seeburg P.H., Adelman J.P.;			
RT	"Characterization of cDNA for precursor of human luteinizing hormone releasing hormone."			
RL	Nature 311:666-668(1984).			
RN	[4]			
RP	SEQUENCE OF 24-33.			
RX	MEDLINE-83126573; Pubmed-6760865;			
RA	Tan L., Rousseau P.;			
RT	"The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in the human placenta."			
RL	Biochem. Biophys. Res. Commun. 109:1061-1071(1982).			
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC	-1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYEST LABS), LUTREPUSE OR LUTRELER (FERRING PHARMACEUTICALS) AND RELISORM (SERONO).			
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: X01059; CAA25526.1; -			
DR	EMBL: M12578; AAA35916.1; -			
DR	EMBL: X15215; CAA33285.1; -			
DR	PIR: A01410; RHUG.			
DR	PIR: A26173; A26173.			
DR	PIR: S05308; S05308.			
DR	MTM: 152760; -			
DR	INTERPRO: IPR002012; -			
DR	PRAM: PF00446; GNRH.1.			
DR	PROSITE: PS00473; GNRH.1.			

KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Pharmaceutical; Signal.

FT	SIGNAL	1	23	
FT <td>CHAIN</td> <td>24</td> <td>92</td> <td>PROGONADOLIBERIN I.</td>	CHAIN	24	92	PROGONADOLIBERIN I.
FT <td>PEPTIDE</td> <td>24</td> <td>33</td> <td>GONADOLIBERIN I.</td>	PEPTIDE	24	33	GONADOLIBERIN I.
FT <td>PEPTIDE</td> <td>37</td> <td>92</td> <td>GNRH-ASSOCIATED PEPTIDE I.</td>	PEPTIDE	37	92	GNRH-ASSOCIATED PEPTIDE I.
FT <td>ACT_SITE</td> <td>26</td> <td>26</td> <td>APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.</td>	ACT_SITE	26	26	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT <td>MOD_RES</td> <td>34</td> <td>24</td> <td>PYRROLIDONE CARBOXYLIC ACID.</td>	MOD_RES	34	24	PYRROLIDONE CARBOXYLIC ACID.
FT <td>MOD_RES</td> <td>33</td> <td>33</td> <td>AMIDATION (G-34 PROVIDE AMIDE GROUP).</td>	MOD_RES	33	33	AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT <td>CONFLICT</td> <td>16</td> <td>16</td> <td>W -&gt; S (IN REF. 3).</td>	CONFLICT	16	16	W -> S (IN REF. 3).
SO	SEQUENCE	92 AA;	10380 MW;	30A72221B076FA79 CRC64;

Query Match Score 76.5; DB 1; Length 92;  
Best Local Similarity 78.96; Pred. No. 0.063;  
Matches 15; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Db 20 GCSSQHSYGLRPG-GRKD 37

RESULT 5

ID	GONI_MOUSE	STANDARD	PRT	90 AA
AC	P13562			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I) (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR I].			
DE	GNRH I OR GNRH.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-87069928; Pubmed-3024317;			
RA	Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,			
RA	Phillips H.S., Nikolics K., Seeburg P.H.;			
RT	"A deletion truncating the gonadotropin-releasing hormone gene is responsible for hypogonadism in the hpg mouse."			
RL	Science 234:1366-1371(1986).			
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: M14872; AAA37717.1; -			
DR	MGD: MGI:95789; GNRH.			
DR	INTERPRO: IPR002012; -			
DR	PRAM: PF00446; GNRH.1.			
DR	PROSITE: PS00473; GNRH.1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Signal.			
FT <td>SIGNAL</td> <td>1</td> <td>21</td> <td></td>	SIGNAL	1	21	
FT <td>CHAIN</td> <td>22</td> <td>90</td> <td>PROGONADOLIBERIN I.</td>	CHAIN	22	90	PROGONADOLIBERIN I.
FT <td>PEPTIDE</td> <td>22</td> <td>31</td> <td>GONADOLIBERIN I.</td>	PEPTIDE	22	31	GONADOLIBERIN I.
FT <td>PEPTIDE</td> <td>35</td> <td>90</td> <td>PROLACTIN RELEASE-INHIBITING FACTOR I.</td>	PEPTIDE	35	90	PROLACTIN RELEASE-INHIBITING FACTOR I.
FT <td>ACT_SITE</td> <td>24</td> <td>24</td> <td>APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.</td>	ACT_SITE	24	24	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT <td>MOD_RES</td> <td>22</td> <td>22</td> <td>PYRROLIDONE CARBOXYLIC ACID.</td>	MOD_RES	22	22	PYRROLIDONE CARBOXYLIC ACID.

FT MOD\_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 90 AA: 10337 MW: 1C076FEFA4826E4D9 CRC64;  
 Query Match 25.5%; Score 74; DB 1; Length 90;  
 Best Local Similarity 92.9%; Pred. No. 0.11;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 23 GGSSQHWSTGLRPG 36  
 Db 18 GCSSQHWSTGLRPG 31

RESULT 6  
 GONL\_PIG  
 ID GONL\_PIG STANDARD; PRT; 91 AA.  
 AC P49921.  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PROGNADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)  
 (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
 HORMONE I) (GNRH I) (LULIBERIN I); GNRH ASSOCIATED PEPTIDE I].  
 GN GNRH OR GNRH.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HYPOTHALAMUS;  
 RA Weesner G.D., Maltieri R.L., Becker B.A.;  
 RL Submitted (May-1994) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 24-33.  
 RX MEDLINE=72114303; PubMed=4946067;  
 RA Baba Y., Matsuo H., Schally A.V.;  
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.  
 RT Confirmation of the proposed structure by conventional sequential  
 RT analyses.";  
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).  
 RN [3]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE=72065376; PubMed=4942726;  
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;  
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-  
 RT phase method.";  
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).  
 RN [4]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE=72117544; PubMed=4946275;  
 RA Baba Y., Arimura A., Schally A.V.;  
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";  
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).  
 CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, J32864; AAA31066.1; -  
 CC PIR, A01411; RHPGG.  
 CC INTERPRO: IPR002012; -  
 CC PEAM: PF00446; GNRH.1.  
 CC PROSITE: PS00473; GNRH.1.  
 KW cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal.

FT SIGNAL 1 23  
 FT CHAIN 24 91 PROGNADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.  
 FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACT\_SITE 26 26 ACTIVITY.  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 91 AA: 10090 MW: 8340474F32DDA99 CRC64;  
 Query Match 25.5%; Score 74; DB 1; Length 91;  
 Best Local Similarity 92.9%; Pred. No. 0.11;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 23 GGSSQHWSTGLRPG 36  
 Db 20 GCSSQHWSTGLRPG 33

RESULT 7  
 GONL\_RAT  
 ID GONL\_RAT STANDARD; PRT; 92 AA.  
 AC P07490.  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PROGNADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)  
 (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING  
 HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR  
 DE I].  
 GN GNRH OR GNRH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86094338; PubMed=2867548;  
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;  
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor  
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
 RT factor in human and rat.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89384661; PubMed=2476669;  
 RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;  
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and  
 RT hypothalamic expression.";  
 RL Mol. Endocrinol. 3:1257-1262(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=THYMUS;  
 RX MEDLINE=93105480; PubMed=1468115;  
 RA Walter C.C., Marchetti B., Lebouef R.D., Blajock J.E.;  
 RT "Thymocytes express a mRNA that is identical to hypothalamic  
 RT luteinizing hormone-releasing hormone mRNA.";  
 RL Cell. Mol. Neurobiol. 12:447-454(1992).  
 RN [4]  
 RP SEQUENCE OF 1-47 FROM N.A.  
 RC TISSUE=HEART;  
 RX MEDLINE=87149087; PubMed=3547652;  
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;  
 RT "Two mammalian genes transcribed from opposite strands of the same  
 RT DNA locus.";  
 RL Science 235:1514-1517(1987).  
 CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -I- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.  
 CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC or send an email to license@lsb.ch.
CC
CC EMBL, J05443; AAA42192.1; -.
CC DR PIR; A36443; A36443.
CC DR INTERPRO; IPRO02080; -.
CC DR PROSITE; PS00515; SVP_II; 12.
CC KW Testosterone; Seminal vesicle; Signal; Repeat; Copulatory plug.
CC FT SIGNAL 1 22
CC FT CHAIN 23 414 SEMINAL VESICLE SECRETORY PROTEIN 11.
CC FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.
CC FT DOMAIN 108 311 13 X 13 AA TANDEM REPEATS.
CC FT REPEAT 108 120 1.
CC FT REPEAT 127 139 2.
CC FT REPEAT 140 152 3.
CC FT REPEAT 153 165 4.
CC FT REPEAT 166 178 5.
CC FT REPEAT 179 191 6.
CC FT REPEAT 192 204 7.
CC FT REPEAT 205 217 8.
CC FT REPEAT 224 236 9.
CC FT REPEAT 237 249 10.
CC FT REPEAT 257 269 11.
CC FT REPEAT 275 287 12.
CC FT REPEAT 299 311 13.
CC SQ SEQUENCE 414 AA; 45539 MW; C89E3FCE0CEE92E CRC64;
CC
CC Query Match 24.5%; Score 71; DB 1; Length 414;
CC Best Local Similarity 42.9%; Pred. No. 0.99;
CC Matches 18; Conservative 3; Mismatches 21; Indels 0; Gaps 0;
CC
CC QY 4 SYGLRPGSGSDMSYGLRPGGSSQHWSYGLRPGSGSDMSYVG 45
CC 1:1 1111 1:1 1:1 1:1 1:1 1:1
CC Db 184 SFGQMKSSGSGVSKSFGQMKKASQGIKSPGQKKSQGGQIQSYVG 225
CC
CC RESULT 9
CC FUS_MOUSE
CC ID FUS_MOUSE STANDARD; PRT; 518 AA.
CC AC P56959;
CC DT 01-OCT-2000 (Rel. 40, Created)
CC DT 01-OCT-2000 (Rel. 40, Last sequence update)
CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
CC DE RNA-BINDING PROTEIN FUS (PIGEPIN PROTEIN).
CC GN FUS.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC RA [1]
CC RP SEQUENCE FROM N.A.
CC RA Alappet S.R., Zhang M., Zhao X., Alliegro M.A., Alliegro M.C.,
CC RA Burdall C.A.;
CC RL "Regulation of pigpen expression in mouse embryos.";
CC RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC CC -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND
CC PROMOTES APP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-
CC STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED
CC DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY
CC SIMILARITY).
CC CC -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY
CC SIMILARITY).
CC CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-sib.ch/announce/
CC or send an email to license@lsb.ch).
CC

```

DR EMBL: AF224264; AAF70602.1; -  
 DR MGD: MGI:1353633; FUS.  
 DR PROSITE: PS50102; RRM: 1.  
 DR RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;  
 DR Metal-binding.  
 KW DOMAIN 1 167 GLN/GLY/SER/TYR-RICH.  
 FT DOMAIN 168 264 GLY-RICH.  
 FT DOMAIN 278 365 RNA-BINDING (RRM).  
 FT DOMAIN 364 518 ARG/GLY-RICH.  
 FT ZN\_FING 421 440 CA-TYPE (POTENTIAL).  
 FT ZN\_FING 421 440  
 SQ SEQUENCE 518 AA; 52673 MW; 306F231BEED78D6 CRC64;

Query Match 24.1%; Score 70; DB 1; Length 518;  
 Best Local Similarity 40.8%; Pred. No. 1.6; Mismatches 15; Indels 12; Gaps 2;  
 Matches 20; Conservative 2; Mismatches 15; Indels 12; Gaps 2;

OY 5 YGLRPGSGSDMSYGLRPGSSQHWMS-----YGLRPGSGSDMSYSG 45  
 Db 101 YGQGPAPSSSTSGSY---GSSSQSSSYGDPQSGGYGQSGYGGQOQSYG 145

RESULT 10  
 FUS\_BOVIN STANDARD; PRT; 512 AA.  
 AC 028009;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE RNA-BINDING PROTEIN FUS (PIGEPEN PROTEIN).  
 OS FUS.  
 GN Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-AORTA;  
 RC MEDLINE-96175600; PubMed-8631501;  
 RA Allegro M.C.; Allegro M.A.;  
 RT "A nuclear protein regulated during the transition from active to  
 RT quiescent phenotype in cultured endothelial cells."  
 RL Dev. Biol. 174:288-297(1996).  
 RN [2]  
 RP CARBOHYDRATE BINDING DOMAIN  
 RX MEDLINE-20160719; PubMed-10694442;  
 RA Allegro M.C.;

RT "A C-terminal carbohydrate-binding domain in the endothelial cell  
 RT regulatory protein, pigpen: new function for an EWS family member."  
 RL Exp. Cell Res. 255:270-277(2000).  
 CC -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND  
 CC PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-  
 CC STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED  
 CC DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY  
 CC SIMILARITY).  
 CC -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR, EXHIBITS DIFFUSE STAINING  
 CC THROUGHOUT (EXCLUDING NUCLEOLI), TOGETHER WITH A SMALL NUMBER OF  
 CC INTENSELY STAINED FOCAL POINTS, OR GRANULES, AND PUNCTATE STAINING  
 CC ALONG THE NUCLEAR ENVELOPE.  
 CC -1- DOMAIN: THE C-TERMINAL DOMAIN BINDS CARBOHYDRATES.  
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
 CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U26024; AAC13543.1; -  
 DR HSSP: P09651; 1HA1.  
 DR INTERPRO: IPR000504; -  
 DR INTERPRO: IPR001876; -  
 DR PFAM: PF00076; trm; 1.  
 DR PFAM: PF00641; zt-RanBP; 1.  
 DR PROSITE: PS50102; RRM; 1.  
 DR PROSITE: PS00030; RRM\_RNP\_1; FALSE NEG.  
 KW RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;  
 KW Metal-binding.  
 FT DOMAIN 1 164 GLN/GLY/SER/TYR-RICH.  
 FT DOMAIN 165 253 GLY-RICH.  
 FT DOMAIN 271 357 RNA-BINDING (RRM).  
 FT DOMAIN 357 512 ARG/GLY-RICH.  
 FT ZN\_FING 414 433 CA-TYPE (POTENTIAL).  
 FT ZN\_FING 414 433  
 SQ SEQUENCE 512 AA; 52240 MW; 3652329C044F1386 CRC64;

Query Match 23.8%; Score 69; DB 1; Length 512;  
 Best Local Similarity 38.0%; Pred. No. 1.9; Mismatches 15; Indels 14; Gaps 2;  
 Matches 19; Conservative 2; Mismatches 15; Indels 14; Gaps 2;

OY 5 YGLRP-----GSGSDMSYGLRPGSSQHWSTYGLRPGSGSDMSYSG 45  
 Db 100 YGQGPAPSSSTSGSYGSSSQSSGYGDPQGC---GYGQSGYGGQOQSYG 144

RESULT 11  
 FUS\_HUMAN STANDARD; PRT; 526 AA.  
 ID FUS\_HUMAN  
 AC P35637;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE RNA-BINDING PROTEIN FUS (ONCOGENE FUS) (ONCOGENE TLS) (TRANSLOCATED IN  
 DE LIPOSARCOMA PROTEIN) (POMF/5) (75 KDA DNA-PAIRING PROTEIN).  
 DE LIPOSARCOMA PROTEIN)  
 GN FUS OR TLS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A. (LONG ISOFORM).  
 RX MEDLINE-93288139; PubMed-8510758;  
 RA Crozat A., Aman P., Mandahl N., Ron D.;  
 RT "Fusion of CHOP to a novel RNA-binding protein in human myxoid  
 RT liposarcoma."  
 RL Nature 363:640-644(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
 RX MEDLINE-93350637; PubMed-7503811;  
 RA Rablitts T.H., Forster A., Larson R., Nathan P.;  
 RT "Fusion of the dominant negative transcription regulator CHOP with a  
 RT novel gene FUS by translocation t(12;16) in malignant liposarcoma."  
 RL Nat. Genet. 4:175-180(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).  
 RX MEDLINE-99013873; PubMed-9795213;  
 RA Morohoshi F., Ootsuka Y., Arai K., Ichikawa H., Mitani S.,  
 RA Munakata N., Ohki M.;  
 RT "Genomic structure of the human RBP56/hRAF168 and FUS/TLS genes."  
 RL Gene 221:191-198(1998).  
 RN [4]  
 RP SEQUENCE OF 235-244; 307-312; 335-345 AND 349-357, & CHARACTERIZATION.  
 RX MEDLINE-20036580; PubMed-10567410;  
 RA Baechtold H., Kuroda M., Sok J., Kon D., Lopez B.S., Akhmedov A.T.;  
 RT "Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein  
 RT TLS/FUS and is able to promote D-loop formation."  
 RL J. Biol. Chem. 274:34337-34342(1999).  
 RN [5]  
 RP CHROMOSOMAL TRANSLOCATION.  
 RX MEDLINE-94243799; PubMed-8187069;  
 RA Ichikawa H., Shimizu K., Hayashi Y., Ohki M.;  
 RT "An RNA-binding protein gene, TLS/FUS, is fused to ERG in human



RT myeloid leukemia with t(16;21) chromosomal translocation.";  
 RL Cancer Res. 54:2865-2868(1994).  
 CC -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND  
 CC PROMOTES ATP-DEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-  
 CC STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED  
 CC DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY.  
 CC -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: UBQUITOUS.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG ISOFORM (SHOWN HERE)  
 CC AND A SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- DISEASE: A FORM OF MALIGNANT MYXOID LIPOSARCOMA IS CHARACTERIZED  
 CC BY A CHROMOSOMAL TRANSLOCATION T(12;16)(Q13;P11) THAT INVOLVES FUS  
 CC AND CHOP.  
 CC -1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED  
 CC BY A CHROMOSOMAL TRANSLOCATION T(16;21)(P11;Q22) THAT INVOLVES  
 CC FUS AND ERG.  
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
 CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.  
 CC -----  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: S62140; AAB27102.1; -;  
 DR EMBL: S62138; AAB27103.1; ALT\_SEQ.  
 DR EMBL: X71427; CA505558.1; ALT\_SEQ.  
 DR EMBL: X71428; CA505559.1; ALT\_SEQ.  
 DR EMBL: AF071213; AAC35285.1; -;  
 DR EMBL: AF071213; AAC35284.1; -;  
 DR PIR: S33799; S33799.  
 DR HSSP: P09651; 1HA1.  
 DR MIM: 137070; -;  
 DR MIM: 151900; -;  
 DR INTERPRO: IPR000504; -;  
 DR INTERPRO: IPR001876; -;  
 DR PFM: PF00076; trm; 1.  
 DR PFM: PF00641; zf-RanBP; 1.  
 DR PROSITE: PS50102; RRM; 1.  
 DR PROSITE: PS00030; RRM\_RNP; 1; FALSE\_NEG.  
 KW Proto-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat;  
 KW Alternative splicing; Chromosomal translocation; Zinc-finger; Zinc;  
 KW Metal-binding.  
 FT DOMAIN 1 165 GLN/GLY/SER/TYR-RICH.  
 FT DOMAIN 166 267 GLY-RICH.  
 FT DOMAIN 285 371 RNA-BINDING (RRM).  
 FT DOMAIN 371 526 ARG/GLY-RICH.  
 FT ZN\_FING 428 447 C4-TYPE (POTENTIAL).  
 FT SITE 266 267 BREAKPOINT FOR TRANSLOCATION TO FORM  
 FT FUS/CHOP ONCOGENE.  
 FT VARSPIC 64 65 TG -> S (IN SHORT ISOFORM).  
 FT CONFLICT 338 338 T -> N (IN REF. 4).  
 SQ SEQUENCE 526 AA; 53426 MW; 88C8E263B7905549 CRC64;

Query Match 23.8%; Score 69; DB 1; Length 526;  
 Best Local Similarity 40.0%; Pred. No. 2;  
 Matches 18; Conservative 2; Mismatches 21; Indels 4; Gaps 1;

Db 5 YGLRPGSGSQMSTYGLRPGSS---QHMSYGLRPGSGSQMSTYG 45  
 100 YGQPPAPSTSTGSSSSSSSSGCGPQSGSYQPPSYGCGQOOSYVG 144

RESULT 12  
 GONI\_HAPBU STANDARD; PRT; 94 AA.  
 ID GONI\_HAPBU  
 AC P51918; O93387;  
 DT 01-OCT-1996 (Ref. 34, Created)

DT 30-MAY-2000 (Ref. 39, Last sequence update)  
 DE 30-MAY-2000 (Ref. 39, Last annotation update)  
 DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)  
 DE (LH-RH I) (LH-RH I).  
 GN GNRH1.  
 OS Haplochromis burtoni.  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;  
 CC Cichlidae; Astatotilapia.  
 CC [1]  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=95396797; PubMed=7667296;  
 RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;  
 RT "Three gonadotropin-releasing hormone genes in one organism suggest  
 RT novel roles for an ancient peptide."  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99061842; PubMed=9843638;  
 RA White R.B., Fernald R.D.;  
 RT "Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression  
 RT reveals a distinct origin for GNRH-containing neurons in the  
 RT midbrain."  
 RL Gen. Comp. Endocrinol. 112:322-329(1998).  
 RN [3]  
 RP SEQUENCE OF 23-32.  
 RC TISSUE=PIUITARY;  
 RX MEDLINE=95372591; PubMed=7644702;  
 RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,  
 RA White S.A., Francis R.C., Fernald R.D., Licht P., Wabry C.,  
 RA Sherwood N.M.;  
 RT "Primary structure of solitary form of gonadotropin-releasing hormone  
 RT (GNRH) in cichlid pituitary; three forms of GNRH in brain of cichlid  
 RT and pumpkinseed fish."  
 RL Regul. Pept. 57:43-53(1995).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE  
 CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALMIC-PITUITARY-  
 CC GONADAL AXIS.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS  
 CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOPHYSAL AXONS.  
 CC -1- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.  
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
 CC -----  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U31865; AAC59691.1; -;  
 DR EMBL: AF076961; AAC27716.1; -;  
 DR INTERPRO: IPR002012; -;  
 DR PFM: PF00446; GNRH; 1.  
 DR PROSITE: PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Multigene family.  
 FT SIGNAL 1 22  
 FT CHAIN 23 94  
 FT PEPTIDE 23 32  
 FT PEPTIDE 36 94  
 FT MOD\_RES 23 32  
 FT MOD\_RES 32 32  
 FT CONFLICT 86 94  
 SQ SEQUENCE 94 AA; 10382 MW; E57DBA83333278D7 CRC64;

Query Match 23.6%; Score 68.5; DB 1; Length 94;  
 Best Local Similarity 65.0%; Pred. No. 0.42;  
 Matches 13; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Db 1 GONADOLIBERIN I.  
 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).  
 PYROLIDONE CARBOXYLIC ACID.  
 AMIDATION (G-33 PROVIDE AMIDE GROUP).  
 ENGRTEPKK -> KMDGSHRNREKFL (IN REF. 1).

```

DE (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
CC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FOREBRAIN;
RX MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Wray S., Batley J.F.;
RT "The frog gonadotropin-releasing hormone-1 (GnRH-1) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GnRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL, L28040; AAA49728.1; -.
CC INTERPRO; IPR02012; -.
CC DR PFAM; PF00446; GnRH; 1.
CC PROSITE; PS00473; GnRH; 1.
CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal.
CC FT SIGNAL 1 23
CC FT CHAIN 24 89
CC FT PEPTIDE 24 33
CC FT CHAIN 37 89
CC FT PEPTIDE 37 85
CC FT MOD_RES 24 24
CC FT MOD_RES 33 33
CC SQ SEQUENCE 89 AA; 10246 MW; 6FAFF36FBAE0D4284 CRC64;
-----
Query Match 22.9%; Score 66.5; DB 1; Length 89;
Best Local Similarity 75.0%; Pred. No. 0.65;
Matches 12; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Oy 26 SOHWYGLRPGSGSOD 41
Db 23 KOHWYGLRPG-GKRD 37

RESULT 15
SD ISOA_PSEAY STANDARD; PRT; 776 AA.
AC P10342.
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ISOAMYLASE PRECURSOR (EC 3.2.1.68).
GN IAM.
OS Pseudomonas amylobacteramosa.
OC Bacteria; Proteobacteria.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SB-15;
RA MEDLINE=86243808; PubMed=3379068;
RA Amemura A., Chakraborty R., Fujita M., Nouni T., Futai M.;
RT "Cloning and nucleotide sequence of the isoamylase gene from
RT Pseudomonas amylobacteramosa SB-15."
RL J. Biol. Chem. 263:9271-9275(1988).
RN [2]
RP SEQUENCE FROM N.A.

```

Search completed: March 2, 2001, 11:01:21  
Job time: 436 sec

RC STRAIN-JD210;  
RX MEDLINE-91064385; Pubmed-2248978;  
RA Chen J.H., Chen Z.Y., Chow T.Y., Chen J.C., Tan S.T., Hsu W.H.;  
RT "Nucleotide sequence and expression of the isoamylase gene from an  
isoamylase-hyperproducing mutant, Pseudomonas amyloclavata JD210.";  
RL Biochim. Biophys. Acta 1087:309-315(1990).  
RN [3]  
RP SEQUENCE OF 744-776 FROM N.A.  
RC STRAIN-SR-15;  
RX MEDLINE-89327147; Pubmed-2753857;  
RA Amemura A., Fujita M., Futai M.;  
RT "Transcription of the isoamylase gene (iam) in Pseudomonas  
amyloclavata SR-15.";  
RL J. Bacteriol. 171:4320-4325(1989).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE-98387895; Pubmed-9719642;  
RA Katsuya Y., Mezaki Y., Kubota M., Matsuura Y.;  
RT "Three-dimensional structure of Pseudomonas isoamylase at 2.2-A  
resolution.";  
RL J. Mol. Biol. 281:885-897(1998).  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,6-ALPHA-D-GLUCOSIDIC BRANCH  
LINKAGES IN GLYCOGEN, AMYLOPECTIN AND THEIR BETA-LIMITS DEXTRINS.  
CC -1- INDUCTION: BY MALTOSE.  
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
KNOWN AS THE ALPHA-AMYLASE FAMILY, ISOAMYLASE SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J03871; AAA25854.1; -;  
DR EMBL; X13378; CAA31754.1; -;  
DR PIR; A28109; A28109.  
DR PDB; 1BF2; 12-AUG-98.  
DR INTERPRO; IPR000461; -;  
DR PRAM; PF00128; alpha-amylase; 1.  
KW Hydrolyase; Glycosidase; Signal; 3D-structure.  
FT SIGNAL 1 26  
FT CHAIN 27 776  
FT ACT\_SITE 401 401  
FT ACT\_SITE 481 481  
FT ACT\_SITE 536 536  
FT DISULFID 410 422  
FT DISULFID 546 616  
FT DISULFID 738 766  
FT CONFLICT 8 8  
FT CONFLICT 126 126  
FT CONFLICT 169 171  
FT CONFLICT 386 386  
FT CONFLICT 413 416  
FT CONFLICT 454 489  
FT CONFLICT 555 556  
FT CONFLICT 650 657  
SO SEQUENCE 776 AA; 83626 MW; F738BF8040246169 CRC64;  
A -> G (IN REF. 1).  
F -> C (IN REF. 1).  
GAS -> AH (IN REF. 1).  
L -> V (IN REF. 1).  
GAYT -> AVH (IN REF. 1).  
SGLDLFAEPWATIGNSYOLGFGPGQWSENGCLFRDS ->  
TWICLRLNLCPSAATRTSWVDSRRVAVENSVPRQ (IN  
REF. 1).  
WP -> S (IN REF. 1).  
AFRAKHPA -> RSARHIP (IN REF. 1).  
Query Match 22.88; Score 66; DB 1; Length 776;  
Best Local Similarity 35.14; Pred. NO. 5.9;  
Matches 13; Conservative 6; Mismatches 16; Indels 2; Gaps 1;  
OY 5 YGLRPGSGSDWGYGLRPGSSOH--WSYGLRPGSGS 39  
DB 544 YSCNANNSQAMPYGPDSGDTSTNYSWDQMSAGTGA 580

**THIS PAGE BLANK (USPTO)**

A;Title: Developmental expression of the genes encoding transforming growth factor  $\alpha$   
A;Reference number: 158134; MIMD:05124501



A:Cross-references: GB:M1670; NID:g204447; PIDN:AAA1264.1; PID:g204448  
R:Adelman, J.P.; Mason, A.J.; Haylick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A>Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotrophin releasing hormone  
A:Reference number: A94090; MUID:86094338  
A:Accession: B26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA11263.1; PID:g204446  
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.  
Cell. Mol. Neurobiol. 12, 447-454, 1992  
A>Title: TLMycocytes express a RNA that is identical to hypothalamic luteinizing hormone releasing hormone  
A:Reference number: A48410; MUID:93105480  
A:Accession: A48410  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-92 <MAI>  
A:Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060  
A:Experimental source: thymus  
A>Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIP:121083)  
C:Genetics:  
A:Introns: 47/3; 79/3  
C:Function:  
A>Description: stimulates pituitary secretion of luteotropin and follitropin  
A>Note: gonadoliberin associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid; n...  
F:1-23/Domin: signal sequence #status predicted <SIG>  
F:24-92/Product: progonadoliberin #status predicted <PGN>  
F:24-33/Product: gonadoliberin #status predicted <GLN>  
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIRF>  
F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following glyx

Query Match            25.5%: Score 74; DB 1; Length 92;  
Best Local Similarity   92.9%: Pred. No. 0.19;  
Matches   13: Conservative   0; Mismatches   1; Indels   0; Gaps   0;

OY         23 GGSSQDMSYGRLPG 36  
            I|||||IIIIII  
DB        20 GCSSQHMSYGLRPG 33

RESULT          6  
G02127  
fus-like protein - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text-change 23-Jul-1999  
C:Accession: G02127  
R:Itoh, K.; Kawase, M.  
submitted to the EMBL Data Library, September 1995  
A:Reference number: G09199  
A:Accession: G02127  
A>Status: preliminary; translated from GB/EMBL/DDBU  
A:Molecule type: mRNA  
A:Residues: 1-528 <ITD>  
A:Cross-references: EMBL:U35561; NTD:g1040969; PIDN:AA79948.1; PID:g1040970  
C:Superfamily: RNA-binding protein EMS; ribonucleoprotein repeat homology  
F:285-364/Domain: ribonucleoprotein repeat homology <RMM>

Query Match            24.8%: Score 72; DB 2; Length 528;  
Best Local Similarity   40.0%: Pred. No. 1.8;  
Matches   20: Conservative   2; Mismatches   12; Indels   16; Gaps   2;

OY         5 YGLRP-----GSQSQDMSYGLRPGSSSQHMSYGLRPGSGSQDMSYG 45  
            ||||     |||||     |     |||||     |||||     |     ||||  
DB        99 YGPGRTPSTSGSYGSSSQTSSYGP-----SYGGQPVGGGQQSGYG 141

RESULT          7  
142382

guanylate cyclase (EC 4.6.1.2) 1, retinal - chicken (fragment)  
N:Alternate names: guanylyl cyclase; retinal guanylate cyclase 1  
C:Species: Gallus gallus (chicken)  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jun-2000  
C:Accession: T42382  
R:Sample-Rowland, S.L.; Lee, N.R.; Van Hooser, J.P.; Palczewski, K.; Baehr, W.  
A:Title: A null mutation in the photoreceptor guanylate cyclase gene causes the retin  
A:Reference number: T42382  
A:Accession: T42382  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1068 <SEM>  
A:Cross-references: EMBL:AF036942; NID:g2828021; PID:g2828022; PID:AA04500.1  
C:Genetics: A:Experimental source: strain Rhode Island Red; retina-pigment epithelium-choroid  
A:Gene: GCL  
C:Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain h  
C:Keywords: cGMP biosynthesis; glycoprotein; phosphorus-oxygen lyase; transmembrane p

Query Match 24.8%; Score 72; DB 2; Length 1068;  
Best Local Similarity 39.0%; Pred. No. 3.8;  
Matches 16; Conservative 3; Mismatches 14; Indels 8; Gaps 2;

QY 10 GSGSQDWS-YGLRPG-----GSSQHWYGLRPGSGSQDW 42  
Db 356 GKGDRLMPVYGLPGIRGLAYRGSHVHPHSSSGTDSGCW 396

RESULT 8  
A36443  
seminal vesicle secretory protein II precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 05-Nov-1999  
C:Accession: A36443  
R:Harris, S.E.; Harris, M.A.; Johnson, C.M.; Bean, M.F.; Dodd, J.G.; Matusik, R.J.; C  
J. Biol. Chem. 265, 9896-9903, 1990  
A:Title: Structural characterization of the rat seminal vesicle secretion II protein  
A:Reference number: A36443; MUID:90277684  
A:Accession: A36443  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-414 <HAR>  
A:Cross-references: GB:J05443; NID:g207114; PID:AAA42192.1; PID:g207115

Query Match 24.5%; Score 71; DB 2; Length 414;  
Best Local Similarity 42.9%; Pred. No. 1.8;  
Matches 18; Conservative 3; Mismatches 21; Indels 0; Gaps 0;

QY 4 SYGLRPGSGSQDWSYGLRPGSSQHWMSYGLRPGSGSQDWSYG 45  
Db 184 SFGQMKSSGSQVSKFSGQMKASESQIKSFQKSSGQGLSYTG 225

RESULT 9  
T16933  
hypothetical protein T24D8.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T16933  
R:Martin, J.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid T24D8.  
A:Reference number: Z18609  
A:Accession: T16933  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-171 <MAR>  
A:Cross-references: EMBL:U04042; NID:g1065529; PID:g1065532; PID:AAA61446.1; CESP:IT2  
A:Genetics: A:Gene: CESP:T24D8.5







**THIS PAGE BLANK (08PT0)**

us-09-306-689-10.rap

/cgn2\_6/  
/cgn2\_6/  
/cgn2\_6/

/cgn2\_6/

```

/cgcn2_6/ptodata/2/paa/US093_COMB.pep:US-09-374-721A-31 + 105.00 235.28 0.0004
/cgcn2_6/ptodata/2/paa/US07_COMB.pep:US-07-761-849-11 + 105.00 235.87 0.0004
/cgcn2_6/ptodata/2/paa/US07_COMB.pep:US-07-761-849-12 + 105.00 235.87 0.0004
/cgcn2_6/ptodata/2/paa/US084_COMB.pep:US-08-476-013-11 + 105.00 235.87 0.0004
/cgcn2_6/ptodata/2/paa/US084_COMB.pep:US-08-476-013-12 + 105.00 235.87 0.0004

seq_name: /cgcn2_6/ptodata/2/paa/US090_COMB.pep:US-09-019-010-04

seq_documentation_block:
: Sequence 4, Application US/09019010
: GENERAL INFORMATION:
:
: APPLICANT: HARLAND, RICHARD
: APPLICANT: MANN, JOHN G.
: APPLICANT: ACRES, STEPHEN D.
: TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
: TITLE OF INVENTION: MOLECULES
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: ROBINS & ASSOCIATES
: STREET: 90 MIDDLEFIELD ROAD, SUITE 200

```

STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```

      FILING DATE: 05-FEB-1998
      CLASSIFICATION: 424
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 60/036,883
      FILING DATE: 05-FEB-1997
      ATTORNEY/AGENT INFORMATION:
      NAME: MCCracken, THOMAS P.
      REGISTRATION NUMBER: 38,548
      REFERENCE/DOCKET NUMBER: 9001-0035
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (650) 325-7812
      TELEFAX: (650) 325-7823
      INFORMATION FOR SEQ ID NO: 4:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 49 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-09-019-010-4

alignment_scores:
      Quality: 290.00      Length: 49
      Ratio: 5.918        Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-09-019-010-4  ..

Align seg 1/1  to: US-09-019-010-4  from: 1  to: 49

1  CAGCATTTGAGCTAGCGCCTGCGCCCTGGCAGCGGTTCTCAAGATTGAG  50
|||||
1  GlnHstPrpSerTyrGlyLeuArgProGlySerGlnAspPrpSe 17
1  CTACGGCTCGGCTCCGGGTGGCTCTAGCCACAGATTGGAGCTACGGCCGTC  100
|||||
17  rTyrGlyLeuArgProGlySerGlnAspPrpSerGlnHstPrpSerTyrGlyLeuA  34
101 GCCCTGGCAGCGGTAGCCACAGATTGGAGCTACGGCCTCGTCGGGGT  147
|||||
34  rGpProGlySerGlySerGlnAspPrpSerTyrGlyLeuArgProGly  49

```

seq\_name: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:US-09-249-447A-7

seq\_documentation\_block:

```
; Sequence 7, Application US/09249447A
; GENERAL INFORMATION:
; APPLICANT: Manns, Jack G.
; TITLE OF INVENTION: PASSIVE IMMUNIZATION AS A TREATMENT FOR HORMONE
; FILE REFERENCE: 9001-0045
; CURRENT APPLICATION NUMBER: US/09/249,447A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/075,637
; PRIOR FILING DATE: 1998-02-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GnRH-2, Fig.
; US-09-249-447A-7
```

alignment\_scores:

Quality:	290.00	Length:	49
Ratio:	5.918	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-306-689-10 x US-09-249-447A-7 ..

Align seg 1/1 to: US-09-249-447A-7 from: 1 to: 49

```
1 CAGCATGGAGCTACGGCTGCGCCCTGCGAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHistrpserTyrglyLeuArgProglySerGlnHistrpserGlnAspTrpse 17
17 rTyrglyLeuArgProglySerGlnHistrpserTyrglyLeuA 34
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTACGGCTGCGGGT 147
|||||
34 rProglySerGlySerGlnAspTrpserTyrglyLeuArgProgly 49
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-305-924-11

seq\_documentation\_block:

```
; Sequence 11, Application US/09305924A
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
; TYPE: PRT
; ORGANISM: GnRH
; US-09-305-924-11
```

alignment\_scores:

Quality:	290.00	Length:	49
----------	--------	---------	----

Ratio:	5.918	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-306-689-10 x US-09-305-924-11 ..

Align seg 1/1 to: US-09-305-924-11 from: 1 to: 49

```
1 CAGCATGGAGCTACGGCTGCGCCCTGCGAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHistrpserTyrglyLeuArgProglySerGlnHistrpserGlnAspTrpse 17
17 rTyrglyLeuArgProglySerGlnHistrpserTyrglyLeuA 34
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTACGGCTGCGGGT 147
|||||
34 rProglySerGlySerGlnAspTrpserTyrglyLeuArgProgly 49
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-306-689-11

seq\_documentation\_block:

```
; Sequence 11, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; FILE REFERENCE: 9001-0047
; CURRENT APPLICATION NUMBER: US/09/306,689B
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/088,024
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; US-09-306-689-11
```

alignment\_scores:

Quality:	290.00	Length:	49
Ratio:	5.918	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-306-689-10 x US-09-306-689-11 ..

Align seg 1/1 to: US-09-306-689-11 from: 1 to: 49

```
1 CAGCATGGAGCTACGGCTGCGCCCTGCGAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHistrpserTyrglyLeuArgProglySerGlnHistrpserGlnAspTrpse 17
17 rTyrglyLeuArgProglySerGlnHistrpserTyrglyLeuA 34
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTACGGCTGCGGGT 147
|||||
34 rProglySerGlySerGlnAspTrpserTyrglyLeuArgProgly 49
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-383-912-4

seq\_documentation\_block:

```
; Sequence 4, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
```

```

seq_name: /cgr2_6/ptcdata/2/paa/US093_COMB.pep:US-09-305-924-13
seq_documentation block:
? Sequence 13, Application US/09305924A
? GENERAL INFORMATION:
? APPLICANT: Jack G. Manns
? APPLICANT: Stephen D. Acres
? APPLICANT: Richard Harland
? TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
? FILE REFERENCE: 9001-0048
? CURRENT APPLICATION NUMBER: US/09/305,924A
? CURRENT FILING DATE: 1999-05-05
? EARLIER APPLICATION NUMBER: US 60/084,217
? EARLIER FILING DATE: 1998-05-05
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 13
? LENGTH: 695
? TYPE: PRT
? ORGANISM: GnRH

```

US-09-305-924-13

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x US-09-305-924-13 ..

Align seg 1/1 to: US-09-305-924-13 from: 1 to: 695

```
1 CAGCATTTGAGCTACGGCTGCGCCCTGCGAGCGGCTTCTCAAGATTGGAG 50
|||||
9 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerGlnHisTrp 25
51 CTACGGCTGCTGCTCCGGTGGCTCTACGACGATTTGAGCTAGCGGCTGC 100
|||||
25 rTyrglyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 42
101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCTGCGGCT 147
|||||
42 rGProGlySerGlySerGlnHisTrpSerTyrGlyLeuArgProGly 57
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-306-689-13

seq\_documentation\_block:

```
; Sequence 13, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; TITLE OF INVENTION: ANIMALS
; FILE REFERENCE: 9001-0047
; CURRENT APPLICATION NUMBER: US/09/306,689B
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/088,024
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: defines a leukotoxin GnrH chimera
US-09-306-689-13
```

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x US-09-306-689-13 ..

Align seg 1/1 to: US-09-306-689-13 from: 1 to: 695

```
1 CAGCATTTGAGCTACGGCTGCGCCCTGCGAGCGGCTTCTCAAGATTGGAG 50
|||||
9 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerGlnHisTrp 25
51 CTACGGCTGCTGCTCCGGTGGCTCTACGACGATTTGAGCTAGCGGCTGC 100
|||||
25 rTyrglyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 42
101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCTGCGGCT 147
|||||
42 rGProGlySerGlySerGlnHisTrpSerTyrGlyLeuArgProGly 57
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-383-912-16

seq\_documentation\_block:

; Sequence 16, Application US/09383912

```
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-16
```

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x US-09-383-912-16 ..

Align seg 1/1 to: US-09-383-912-16 from: 1 to: 699

```
1 CAGCATTTGAGCTACGGCTGCGCCCTGCGAGCGGCTTCTCAAGATTGGAG 50
|||||
9 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerGlnHisTrp 25
51 CTACGGCTGCTGCTCCGGTGGCTCTACGACGATTTGAGCTAGCGGCTGC 100
|||||
25 rTyrglyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 42
101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCTGCGGCT 147
|||||
42 rGProGlySerGlySerGlnHisTrpSerTyrGlyLeuArgProGly 57
```

seq\_name: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:US-09-383-912-8

seq\_documentation\_block:

```
; Sequence 8, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
```

```

: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: REED & ROBINS LLP
: STREET: 285 HAMILTON AVENUE, SUITE 200
: CITY: PALO ALTO
: STATE: CA
: COUNTRY: USA
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/383,912
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/694,865
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: MCCracken, THOMAS P.
: REGISTRATION NUMBER: 38,548
: REFERENCE/DOCKET NUMBER: 9001-0016.22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)327-3400
: TELEFAX: (415)327-3231
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 977 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-383-912-8

```

```

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment\_block:

US-09-306-689-10 x US-09-383-912-8 ..

Align seg 1/1 to: US-09-383-912-8 from: 1 to: 977

```

1 CAGCATTTGAGCTACGGCCCTGGCCCTGGACGGCTTCTCAAGATTGAG 50
|||||
927 GlnHSTPserTyrGlyLeuArpProGlySerGlnAspTrpse 943
51 CTACGGCCCTGCGTCCGGGTGGCTCTAGCCAGCATTTGGAGCTACGGCTGC 100
|||||
943 rTyGlyLeuArpProGlySerGlnHSTPserTyrGlyLeuA 960
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTACGGCTGCGCGGT 147
|||||
960 rGProGlySerGlySerGlnAspTrpserTyrGlyLeuArpProGly 975

seq_name: /cgn2_6/plodata/2/paa/PCTUS_COMB.pep:PCT-US00-22121-35

seq_documentation_block:
: Sequence 35, Application PC/TUS0022121
: GENERAL INFORMATION:
: APPLICANT: Proteinix Company
: TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
: FILE REFERENCE: IGN-2004WO
: CURRENT APPLICATION NUMBER: PCT/US00/22121
: CURRENT FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 09/026,276
: PRIOR FILING DATE: 1998-02-19
: PRIOR APPLICATION NUMBER: US 09/374,721
: PRIOR FILING DATE: 1999-08-13
: NUMBER OF SEQ ID NOS: 35

```

```

: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 35
: LENGTH: 40
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: polypeptide
: OTHER INFORMATION: antigen
PCT-US00-22121-35

```

```

alignment_scores:
  Quality: 199.50      Length: 49
  Ratio: 5.250        Gaps: 3
  Percent Similarity: 77.551  Percent Identity: 77.551

```

alignment\_block:

US-09-306-689-10 x PCT-US00-22121-35 ..

Align seg 1/1 to: PCT-US00-22121-35 from: 1 to: 40

```

1 CAGCATTTGAGCTACGGCCCTGGCCCTGGACGGCTTCTCAAGATTGAG 50
|||||
1 GlnHSTPserTyrGlyLeuArpProGly.....GlnHSTPse 14
51 CTACGGCCCTGCGTCCGGGTGGCTCTAGCCAGCATTTGGAGCTACGGCTGC 100
|||||
14 rTyGlyLeuArpProGly.....GlnHSTPserTyrGlyLeuA 28
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTACGGCTGCGCGGT 147
|||||
28 rGProGly.....GlnHSTPserTyrGlyLeuArpProGly 40

```

seq\_name: /cgn2\_6/plodata/2/paa/US090\_COMB.pep:US-09-026-276-35

seq\_documentation\_block:

```

: Sequence 35, Application US/09026276
: GENERAL INFORMATION:
: APPLICANT: Kenten, John H
: APPLICANT: Tramonitano, Alfonso
: APPLICANT: Pilon, Aprile L
: APPLICANT: Lohnas, Gerald L
: APPLICANT: Roberts, Steven F
: TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
: FILE REFERENCE: U.S. Patent Application No. 09/026,276
: CURRENT APPLICATION NUMBER: US/09/026,276
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 35
: LENGTH: 40
: TYPE: PRT
: ORGANISM: Porcine
: US-09-026-276-35

```

```

alignment_scores:
  Quality: 199.50      Length: 49
  Ratio: 5.250        Gaps: 3
  Percent Similarity: 77.551  Percent Identity: 77.551

```

alignment\_block:

US-09-306-689-10 x US-09-026-276-35 ..

Align seg 1/1 to: US-09-026-276-35 from: 1 to: 40

```

1 CAGCATTTGAGCTACGGCCCTGGCCCTGGACGGCTTCTCAAGATTGAG 50
|||||
1 GlnHSTPserTyrGlyLeuArpProGly.....GlnHSTPse 14
51 CTACGGCCCTGCGTCCGGGTGGCTCTAGCCAGCATTTGGAGCTACGGCTGC 100
|||||
14 rTyGlyLeuArpProGly.....GlnHSTPserTyrGlyLeuA 28

```

```

101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCTCCGGGT 147
|||||
28 rprpogly.....GlnH1StrpSerTyrglyLeuArpProgly 40

```

seq\_name: /cgn2.6/ptodata/2/paa/US093\_COMB.pep:US-09-374-721A-35

seq\_documentation\_block:

```

; Sequence 35, Application US/09374721A
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H.
; APPLICANT: Roberts, Steven
; APPLICANT: Lohnas, Gerald
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: CIP OF IGN-9601
; CURRENT APPLICATION NUMBER: US/09/374,721A
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 09/026,276
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 35
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; US-09-374-721A-35

```

```

alignment_scores:
  Quality: 199.50      Length: 49
  Ratio: 5.250         Gaps: 3
  Percent Similarity: 77.551  Percent Identity: 77.551

```

alignment\_block:

```

US-09-306-689-10 x US-09-374-721A-35  ..
Align seg 1/1 to: US-09-374-721A-35 from: 1 to: 40

```

```

1 CAGCATTTGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
1 GlnH1StrpSerTyrglyLeuArpProgly.....GlnH1Strpse 14
51 CTACGGCTGCGTCCGGGTGCTCTACCCAGCATTTGAGCTACGGCTGC 100
|||||
14 rTyrglyLeuArpProgly.....GlnH1StrpSerTyrglyLeuArpProgly 28
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCTCCGGGT 147
|||||
28 rprpogly.....GlnH1StrpSerTyrglyLeuArpProgly 40

```

seq\_name: /cgn2.6/ptodata/2/paa/PCTUS\_COMB.pep:PCT-US00-22121-34

seq\_documentation\_block:

```

; Sequence 34, Application PC/TUS0022121
; GENERAL INFORMATION:
; APPLICANT: Proteolinx Company
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: IGN-2004WO
; CURRENT APPLICATION NUMBER: PCT/US00/22121
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 09/026,276
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 09/374,721
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 34
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; OTHER INFORMATION: antigen
PCT-US00-22121-34

```

```

alignment_scores:
  Quality: 199.50      Length: 49
  Ratio: 5.250         Gaps: 3
  Percent Similarity: 77.551  Percent Identity: 77.551

```

alignment\_block:

```

US-09-306-689-10 x PCT-US00-22121-34  ..
Align seg 1/1 to: PCT-US00-22121-34 from: 1 to: 41

```

```

1 CAGCATTTGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
1 GlnH1StrpSerTyrglyLeuArpProgly.....GlnH1Strpse 14
51 CTACGGCTGCGTCCGGGTGCTCTACCCAGCATTTGAGCTACGGCTGC 100
|||||
14 rTyrglyLeuArpProgly.....GlnH1StrpSerTyrglyLeuArpProgly 28
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCTCCGGGT 147
|||||
28 rprpogly.....GlnH1StrpSerTyrglyLeuArpProgly 40

```

seq\_name: /cgn2.6/ptodata/2/paa/US090\_COMB.pep:US-09-026-276-34

seq\_documentation\_block:

```

; Sequence 34, Application US/09026276
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H.
; APPLICANT: Tramoniano, Alfonso
; APPLICANT: Pilon, April L.
; APPLICANT: Lohnas, Gerald F.
; APPLICANT: Roberts, Steven F.
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: U.S. Patent Application No. 09/026,276
; CURRENT APPLICATION NUMBER: US/09/026,276
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 34
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Porcine
; US-09-026-276-34

```

```

alignment_scores:
  Quality: 199.50      Length: 49
  Ratio: 5.250         Gaps: 3
  Percent Similarity: 77.551  Percent Identity: 77.551

```

alignment\_block:

```

US-09-306-689-10 x US-09-026-276-34  ..
Align seg 1/1 to: US-09-026-276-34 from: 1 to: 41

```

```

1 CAGCATTTGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
1 GlnH1StrpSerTyrglyLeuArpProgly.....GlnH1Strpse 14
51 CTACGGCTGCGTCCGGGTGCTCTACCCAGCATTTGAGCTACGGCTGC 100
|||||
14 rTyrglyLeuArpProgly.....GlnH1StrpSerTyrglyLeuArpProgly 28
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCTCCGGGT 147
|||||
28 rprpogly.....GlnH1StrpSerTyrglyLeuArpProgly 40

```





XX 24-JAN-1996; 96WO-CA00049.  
 PF  
 XX  
 PR 10-FEB-1995; 95US-0387156.  
 XX  
 PA (UYSA-) UNIV SASKATCHEWAN.  
 PI  
 PI Manns JG, Potter AA;  
 DR WPI: 1996-184447/38.  
 DR N-PSDB; T37178.  
 XX  
 Gonadotropin-releasing hormone multimer fusion proteins - with  
 PT leukotoxin polypeptide for increased immunogenicity, useful in  
 PT antifertility vaccine prodn.  
 PS  
 PS Example 2; Fig 1B; 87pp; English.  
 XX  
 CC A synthetic DNA sequence (T37178) codes for a gonadotropin  
 CC releasing hormone (GNRH) tetramer (W03944), in which the 4 GNRH  
 CC repeat units are separated by spacers designed to increase  
 CC immunogenicity. The DNA sequence was incorporated into vector  
 CC pAA332 (ATCC 68283), which contains a truncated leukotoxin  
 CC gene (LKT 352) derived from Pasteurella haemolytica, to  
 CC give plasmid PCB113 (T37176). Escherichia coli transformants  
 CC produce an LKT-GNRH fusion protein (see also W03942 and W03943)  
 CC useful as a vaccine for fertility control of domestic or farm animals.  
 CC  
 SQ Sequence 49 AA;  
 XX  
 Query Match 100.0%; Score 290; DB 17; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 49  
 DB 1 qhwsyglrpgsgsdwsyglrpgssqhwsyglrpgsgsdwsyglrpg 49  
 XX  
 RESULT 2  
 W79567  
 ID W79567 standard; Protein; 49 AA.  
 XX  
 AC W79567;  
 XX  
 DT 24-DEC-1998 (first entry)  
 XX  
 DE GNRH-2.  
 XX  
 KM Gonadotropin releasing hormone; GNRH; decapeptide; hypothalamus; spacer;  
 KM LH; luteinising hormone; FSH; follicle stimulating hormone; vertebrate;  
 KM pyroglu; chimera; leukotoxin polypeptide; multimer; vaccine; tumour;  
 KM immunogenic.  
 XX  
 OS Synthetic.  
 XX  
 PN W09806848-A1.  
 XX  
 PD 19-FEB-1998.  
 XX  
 PF 08-AUG-1997; 97WO-CA00559.  
 XX  
 PR 09-AUG-1996; 96US-0694865.  
 XX  
 PA (UYSA-) UNIV SASKATCHEWAN.  
 PI  
 PI Manns JG, Potter AA;  
 XX  
 DR WPI: 1998-159540/14.  
 DR N-PSDB; V61529.  
 XX  
 PT Chimeric protein of leukotoxin and gonadotropin releasing hormone

PT useful for, e.g. preparation of vaccines for reduction of incidence  
 PT of mammary tumours in mammals  
 XX  
 PS Disclosure; Figure 1B; 118pp; English.  
 XX  
 CC The present sequence represents a recombinantly produced or chemically  
 CC synthesised gonadotropin releasing hormone-2 (GNRH-2) polypeptide, which  
 CC contains four copies of the GNRH decapeptide (V61528) and triplet amino  
 CC acid spacers between each of these sequences. This decapeptide is  
 CC secreted naturally by the hypothalamus which controls release of both the  
 CC luteinising hormone (LH) and the follicle stimulating hormone (FSH) in  
 CC vertebrates. This sequence, as compared to the native peptide, has been  
 CC found to have an N-terminal Gln rather than a pyroglu residue, and also  
 CC contains substitutions at amino acid residues 15 and 41, whereby His is  
 CC replaced by Asp. This produces an alternating multimeric GNRH sequence  
 CC which is highly immunogenic that can be used in the construction of a  
 CC chimeric protein that comprises a leukotoxin polypeptide, several  
 CC multimers, and the GNRH sequence. The chimeric protein can be used as a  
 CC vaccine to help reduce the incidence of mammary tumours in a mammalian  
 CC individual.  
 CC  
 SQ Sequence 49 AA;  
 XX  
 Query Match 100.0%; Score 290; DB 19; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 49  
 DB 1 qhwsyglrpgsgsdwsyglrpgssqhwsyglrpgsgsdwsyglrpg 49  
 XX  
 RESULT 3  
 W61542  
 ID W61542 standard; Protein; 49 AA.  
 XX  
 AC W61542;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE Peptide hormone GNRH-2 decapeptide (4 copies) fragment.  
 XX  
 KM GNRH; gonadotropin releasing hormone peptide hormone; leukotoxin;  
 KM immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;  
 KM immune response; hormone receptor; cancerous cell; domestic animal;  
 KM porcine; bovine; luteinizing hormone; follicle stimulating hormone;  
 KM immunocastate.  
 XX  
 OS Synthetic.  
 XX  
 PN W09834639-A1.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PF 04-FEB-1998; 98WO-CA00059.  
 XX  
 PR 05-FEB-1997; 97US-0036883.  
 XX  
 PA (BIOS-) BIOSSTAR INC.  
 PI  
 PI Acres SD, Harland R, Manns JG;  
 DR WPI: 1998-446952/38.  
 DR N-PSDB; V45190.  
 XX  
 PT Immunisation against endogenous molecules by administering vaccine  
 PT to ear - useful to elicit efficient and uniform immune response  
 PT against e.g. gonadotropin releasing hormone to immunocastate pigs  
 PT and cattle  
 XX  
 PS Example 1; Fig 1B; 61pp; English.  
 XX

CC This represents the amino acid sequence of the gonadotropin releasing  
 CC hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-  
 CC GnRH polypeptide gene fusions. This is used to exemplify the method of  
 CC invention of immunisation against endogenous molecules by administering  
 CC a vaccine which comprises an immunogen and a carrier to the ear of the  
 CC mammal. The method is useful for eliciting an efficient and uniform  
 CC immune response to block or suppress the activity of an endogenous  
 CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,  
 CC or to elicit an immune response against a targeted endogenous cell type  
 CC (e.g. a cancerous or otherwise diseased cell). It is especially useful  
 CC to reduce the levels of GnRH in domestic animals, especially in porcine  
 CC or bovine species. The use of GnRH immunogens in the vaccine reduces the  
 CC levels of luteinizing hormone and follicle stimulating hormone and helps  
 CC in immunocastrating the animal. Administration of vaccine compositions to  
 CC the ear instead of intramuscular administration into the neck increases  
 CC the efficiency of vaccination of mammals against endogenous immunogens,  
 CC and may increase uniformity of vaccine presentation since the ear is  
 CC relatively uniform from animal to animal.

XX Sequence 49 AA:

Query Match 100.0%; Score 290; DB 19; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSDMSYGLRPG 49  
 DB 1 qhmsyglrpgsgsdmsyglrpgsgsqhmsyglrpgsgsdmsyglrpg 49

RESULT 4  
 ID Y58363 standard; Protein; 49 AA.

XX Y58363:

DT 27-MAR-2000 (first entry)

XX Four-copy gonadotropin-releasing hormone (GnRH) multimer.

XX GnRH multimer; gonadotropin-releasing hormone; immunosterilisation;

KW immunocastration; vaccine; feline; canine; equine; cervine; ds

XX Mammalia.

OS Synthetic.

XX MO9962545-A2.

PN 09-DEC-1999.

PD 28-MAY-1999; 99WO-CA00493.

XX 04-JUN-1998; 98US-0088024.

PR 06-MAY-1999; 99US-0306689.

XX (BIOS-) BIOSTAR INC.

PI Robbins SC;

XX WPI: 2000-086857/07.

DR N-PSDB: 255702.

XX Hormone immunogens, analogues or antibodies used to manufacture

PT vaccines for suppression of reproductive behavior and fertility in

XX vertebrates -

PS Claim 5; Fig 5B; 88pp; English.

XX This sequence represents a four-copy gonadotropin-releasing

CC hormone (GnRH) multimer, where the second and fourth GnRH sequence

CC have a His to Asp substitution at position 2 of the GnRH sequence.

CC The invention relates to GnRH immunogens, analogues or antibodies

CC that cross-react with endogenous GnRH of a vertebrate. A specifically  
 CC claimed immunogenic fusion protein (Y58361) comprises, in the N to  
 CC C-terminal direction, a synthetic peptide sequence (Y58364), an eight  
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer  
 CC sequence of Y58363), the LKT protein (which functions as a carrier  
 CC protein), and a second eight copy GnRH multimer. The fusion protein may  
 CC be used in a vaccine composition for prepubertal administration to a  
 CC vertebrate subject to result in prolonged suppression of reproductive  
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are  
 CC used to manufacture a composition or vaccine for immunosterilisation or  
 CC immunocastration of feline, canine, equine or cervine subjects.  
 CC The vaccines are used to suppress reproductive behaviour and/or  
 CC fertility for at least 10 months. The prepubertal administration  
 CC results in a prolonged, long-term suppression of testicular development  
 CC and/or function in males, or a prolonged, long-term suppression of  
 CC ovarian development and/or function in females. The methods provide a  
 CC viable and desirable alternative to surgical forms of sterilisation that  
 CC are currently used.

XX Sequence 49 AA:

Query Match 100.0%; Score 290; DB 21; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSDMSYGLRPG 49  
 DB 1 qhmsyglrpgsgsdmsyglrpgsgsqhmsyglrpgsgsdmsyglrpg 49

RESULT 5  
 ID Y58135 standard; Protein; 49 AA.

XX Y58135:

DT 07-MAR-2000 (first entry)

XX GnRH analogue multimer, containing four copies of the GnRH analogue.

XX Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;

KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;

KW non-androgenic; steroid; reduction; weight gain; muscle distribution;

KW fat distribution; male pattern; boar taint; flavour; impairment;

XX reliable; immunocastration; meat production.

XX Synthetic.

XX Mammalia.

XX MO9956771-A2.

PN 11-NOV-1999.

PD 05-MAY-1999; 99WO-CA00360.

XX 05-MAY-1998; 98US-0084217.

XX (BIOS-) BIOSTAR INC.

XX Manns JG, Acres SD, Harland R;

XX WPI: 2000-062125/05.

DR N-PSDB: 246402.

XX Production of uncastrated male food animals using vaccines -

XX Example 1; Fig 2B; 87pp; English.

XX This sequence represents four copies of a gonadotropin

CC releasing hormone (GnRH) analogue, DNA encoding which was

CC used in the construction of a chimeric GnRH-leukotoxin (LKT)

CC fusion gene (Z46400). This fusion gene encodes a GnRH-LKT fusion

CC protein which may be used as a vaccine. The LKT portion of the protein  
 CC acts to enhance the immunogenicity of the GnRH portion. The invention  
 CC relates to a method of using two GnRH immunogen vaccines to produce  
 CC uncastrated male animals for meat production, one vaccination prior to  
 CC or during the fattening period to reduce circulating testosterone levels,  
 CC and the second vaccination about 2-8 weeks before slaughter to  
 CC substantially reduce androgenic and/or non-androgenic steroids. The  
 CC invention is used to produce food animals that exhibit the weight gain  
 CC and muscle/fat distribution of male animals without the problems  
 CC associated with male animals. Such problems include "boar taint", a  
 CC urine-like odour found in cooked meat of uncastrated pigs which is  
 CC caused by steroids stored in the tissues, and similar flavour  
 CC impairments in the meat of other intact male animals. The invention is  
 CC more reliable than prior art immunocastration techniques.

SO Sequence 49 AA;

Query Match 100.0%; Score 290; DB 21; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSDMSYGLRPG 49  
 DB 1 qhwsyglrpgsgsdmsyglrpgssqhmsyglrpgsgsdmsyglrpg 49

RESULT 6

W03943 W03943 standard; Protein; 544 AA.

AC W03943;

DT 20-NOV-1996 (first entry)

DE LKT-GnRH protein fusion from PCB111.

XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;

KW fusion protein; immunogen; vaccine; fertility control;

XX contraceptive; sterilisation.

OS Chimeric Pasteurella haemolytica A1 strain B122;

XX Chimeric synthetic.

FT Key Location/Qualifiers

FT Domain 1..493

FT Domain /label= LKT

FT Domain 494..544

FT Domain /label= GnRH\_repeat\_domain

PN W09624675-A1.

XX 15-AUG-1996.

XX 24-JAN-1996; 96WO-CA00049.

XX 10-FEB-1995; 95US-0387156.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1996-384447/38.

XX N-PSDB: T37177.

XX Gonadotropin-releasing hormone multimer fusion proteins - with

XX leukotoxin polypeptide for increased immunogenicity, useful in

XX antiferility vaccine prodn.

CC and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat  
 CC sequence (see also W03944). It is the product of a chimeric  
 CC gene (T37177) produced by deleting an approx. 1300 bp sequence  
 CC from PCB113 (see also T37176) coding for amino acids 352-784  
 CC of LKT-352. Recombinant plasmid PCB111 (LKT 111:4 copy GnRH,  
 CC ATCC 69748) was obtd. Escherichia coli transformants produced  
 CC the chimeric protein, which is useful as a vaccine for fertility  
 CC control, esp. immunological sterilisation of domestic or farm  
 CC animals.

SO Sequence 544 AA;

Query Match 100.0%; Score 290; DB 17; Length 544;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-24;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSDMSYGLRPG 49  
 DB 494 qhwsyglrpgsgsdmsyglrpgssqhmsyglrpgsgsdmsyglrpg 542

RESULT 7

W79570 W79570 standard; Protein; 544 AA.

AC W79570;

DT 24-DEC-1998 (first entry)

DE LKT-GnRH chimeric protein.

XX Chimera; PCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;

KW cytotoxic activity; antigen presentation; immune response; vaccine;

XX tumour.

XX Synthetic.

PN W09806848-A1.

XX 19-FEB-1998.

XX 08-AUG-1997; 97WO-CA00559.

XX 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1998-159540/14.

XX N-PSDB: V61532.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone

XX useful for, e.g. preparation of vaccines for reduction of incidence

XX of mammary tumours in mammals

XX Disclosure; Figure 7.1-5; 118pp; English.

XX The present sequence represents the LKT-GnRH chimeric protein from

XX PCB111. This plasmid contains the LKT 111 polypeptide fused to

XX four copies of the GnRH peptide. This chimera lacks cytotoxic activity

XX which enables there to be an increase in antigen presentation and thus an

XX optimal immune response. The removal of this region also enables the

XX truncated LKT to be expressed at much higher levels and allows the amount

XX of antigen administered to be reduced. This chimeric protein comprises a

XX leukotoxin polypeptide, several multimers, and a GnRH sequence. The

CC chimeric protein can be used as a vaccine to help reduce the incidence of  
 CC mammary tumours in a mammalian individual.

SO Sequence 544 AA;

Query Match 100.0%; Score 290; DB 17; Length 544;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-24;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSDMSYGLRPG 49  
 DB 494 qhwsyglrpgsgsdmsyglrpgssqhmsyglrpgsgsdmsyglrpg 542

RESULT 7

W79570 W79570 standard; Protein; 544 AA.

AC W79570;

DT 24-DEC-1998 (first entry)

DE LKT-GnRH chimeric protein.

XX Chimera; PCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;

KW cytotoxic activity; antigen presentation; immune response; vaccine;

XX tumour.

XX Synthetic.

PN W09806848-A1.

XX 19-FEB-1998.

XX 08-AUG-1997; 97WO-CA00559.

XX 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1998-159540/14.

XX N-PSDB: V61532.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone

XX useful for, e.g. preparation of vaccines for reduction of incidence

XX of mammary tumours in mammals

XX Disclosure; Figure 7.1-5; 118pp; English.

XX The present sequence represents the LKT-GnRH chimeric protein from

XX PCB111. This plasmid contains the LKT 111 polypeptide fused to

XX four copies of the GnRH peptide. This chimera lacks cytotoxic activity

XX which enables there to be an increase in antigen presentation and thus an

XX optimal immune response. The removal of this region also enables the

XX truncated LKT to be expressed at much higher levels and allows the amount

XX of antigen administered to be reduced. This chimeric protein comprises a

XX leukotoxin polypeptide, several multimers, and a GnRH sequence. The

Query Match 100.0%; Score 290; DB 19; Length 544;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-24;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSGSDMSYGLRPG 49  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 494 qhwsyglrpgsgsqdwsyglrpgsgsqhwsyglrpgsgsqdwsyglrpg 542

## RESULT 8

W79573  
 ID W79573 standard; Protein; 695 AA.

W79573;  
 AC

24-DEC-1998 (first entry)  
 DT

LKT-GnRH chimeric protein.  
 DE

Chimera: PCB122; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;  
 KW cytotoxic activity; antigen presentation; Immune response; vaccine;

tumour.  
 KM

Synthetic.  
 OS

MO9806848-A1.  
 PN

19-FEB-1998.  
 PD

08-AUG-1997; 97WO-CA00559.  
 PF

09-AUG-1996; 96US-0694865.  
 PR

(UYSA-) UNIV SASKATCHEWAN.  
 RA

Manns JG, Potter AA;  
 PI

WPI: 1998-159540/14.  
 DR

N-PSDB: V61535  
 DR

Chimeric protein of leukotoxin and gonadotropin releasing hormone  
 PT useful for, e.g. preparation of vaccines for reduction of incidence

of mammary tumours in mammals  
 PT

Claim 9; Figure 9.1-6; 11pp; English.  
 PS

The present sequence represents the LKT-GnRH chimeric protein from  
 CC PCB122. This plasmid contains the LKT 111 polypeptide fused to sixteen

copies of the GnRH peptide. In the pattern of: 8 copies of GnRH-LKT 111-8  
 CC copies of GnRH. This chimera lacks cytotoxic activity which enables

there to be an increase in antigen presentation and thus an optimal  
 CC immune response. The removal of this region also enables the truncated

LKT to be expressed at much higher levels and allows the amount of  
 CC antigen administered to be reduced. This chimeric protein comprises a

leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
 CC chimeric protein can be used as a vaccine to help reduce the incidence of

mammary tumours in a mammalian individual.  
 CC

Sequence 695 AA;  
 SO

Query Match 100.0%; Score 290; DB 19; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-24;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSGSDMSYGLRPG 49  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 9 qhwsyglrpgsgsqdwsyglrpgsgsqhwsyglrpgsgsqdwsyglrpg 57

RESULT 9  
 Y58361  
 ID Y58361 standard; Protein; 695 AA.

XX Y58361;  
 AC

27-MAR-2000 (first entry)  
 DT

Leukotoxin/gonadotropin-releasing hormone fusion protein.  
 DE

Leukotoxin: gonadotropin-releasing hormone; GnRH; immunostimulation;  
 KW immunosuppression; vaccine; feline; canine; equine; cervine.

Chimeric - Pasteurella haemolytica.  
 OS

Chimeric - Mammalia.  
 OS

MO9962545-A2.  
 PN

09-DEC-1999.  
 PD

28-MAY-1999; 99WO-CA00493.  
 PF

04-JUN-1998; 98US-0088024.  
 PR

06-MAY-1999; 99US-0306689.  
 PR

(BIO-) BIOSTAR INC.  
 PA

Robbins SC;  
 PI

WPI: 2000-086857/07.  
 DR

N-PSDB: Z55700.  
 DR

Hormone immunogens, analogues or antibodies used to manufacture  
 PT vaccines for suppression of reproductive behavior and fertility in

vertebrates -  
 PT

Claim 20; Fig 6A-6F; 88pp; English.  
 PS

This sequence represents a fusion protein comprising gonadotropin-  
 CC releasing hormone (GnRH) immunogens and a Pasteurella

haemolytica leukotoxin (LKT) protein. The fusion protein comprises, in  
 CC the N to C-terminal direction, a synthetic peptide sequence (Y58364), an

eight copy GnRH multimer (composed of two copies of the 4xGnRH multimer  
 CC sequence of Y58363), the LKT protein (which functions as a carrier

protein), and a second eight copy GnRH multimer. The fusion protein may  
 CC be used in a vaccine composition for prepubertal administration to a

vertebrate subject to result in prolonged suppression of reproductive  
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies

that cross-react with endogenous GnRH of a vertebrate subject are used  
 CC to manufacture a composition or vaccine for immunosuppression or

immunosuppression of feline, canine, equine or cervine subjects.  
 CC The vaccines are used to suppress reproductive behaviour and/or

fertility for at least 10 months. The prepubertal administration  
 CC results in a prolonged, long-term suppression of testicular development

and/or function in males, or a prolonged, long-term suppression of  
 CC ovarian development and/or function in females. The methods provide a

viable and desirable alternative to surgical forms of sterilisation that  
 CC are currently used.  
 CC

Sequence 695 AA;  
 SO

Query Match 100.0%; Score 290; DB 21; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-24;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSGSDMSYGLRPG 49  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 9 qhwsyglrpgsgsqdwsyglrpgsgsqhwsyglrpgsgsqdwsyglrpg 57

RESULT 10  
 Y58133  
 ID Y58133 standard; Protein; 695 AA.  
 AC Y58133;



PM WO9806848-A1.  
 XX 19-FEB-1998.  
 PD 08-AUG-1997; 97WO-CA00559.  
 XX 09-AUG-1996; 96US-0694865.  
 XX (UUSA-) UNIV SASKATCHEWAN.  
 PA  
 PI Manns JG, Potter AA:  
 DR WPI: 1998-159540/14.  
 DR N-PSDB; V61531.  
 XX  
 PT Chimeric protein of leukotoxin and gonadotropin releasing hormone  
 PT useful for, e.g. preparation of vaccines for reduction of incidence  
 PT of mammary tumours in mammals  
 PS  
 PS Disclosure: Figure 5.1-8; 118pp; English.  
 XX  
 CC The present sequence represents the LKT-GnRH chimeric protein from  
 CC pCbl13. This plasmid contains the LKT 352 polypeptide (W79568) fused to  
 CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity  
 CC which enables there to be an increase in antigen presentation and thus an  
 CC optimal immune response. The removal of this region also enables the  
 CC truncated LKT to be expressed at much higher levels and allows the amount  
 CC of antigen administered to be reduced. This chimeric protein comprises a  
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
 CC chimeric protein can be used as a vaccine to help reduce the incidence of  
 CC mammary tumours in a mammalian individual.  
 CC  
 XX Sequence 977 AA:  
 SQ  
 Query Match 100.0%; Score 290; DB 19; Length 977;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-24;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 OHWSYGLRPGSGQDWSYGLRPGSSQHMWSYGLRPGSGQDWSYGLRPG 49  
 Db 927 qhwsyglrpgsgqdwmsyglrpgsgsqhmsyglrpgsgsqdwsgyglrpg 975  
 RESULT 13  
 ID Y31183 standard; peptide; 40 AA.  
 XX Y31183:  
 AC 28-OCT-1999 (first entry)  
 DT  
 DE Ubiquitin fusion protein GnRH fragment 2.  
 XX  
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;  
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;  
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;  
 KW pulmonary infection; gamete maturation; prostate; anti-self; pig;  
 KW steriodogenesis; gamete maturation; prostate; anti-self; pig;  
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;  
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;  
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9942472-A1.  
 XX  
 PD 26-AUG-1999.  
 XX  
 PF 26-JAN-1999; 99WO-US01588.  
 XX  
 PR 19-FEB-1998; 98US-0026276.  
 XX

PA (IGEN-) IGEN INT INC.  
 XX  
 PI Kenten JH, Lohmas GL, Pilon AL, Roberts SF, Tramontano A;  
 XX WPI: 1999-518582/43.  
 DR  
 XX  
 PT Epitope-containing fusion proteins used to generate a highly  
 PT specific immune responses  
 PS  
 PS Claim 83; Page 43; 67pp; English.  
 XX  
 CC This invention describes a novel fusion protein, comprising a heat shock  
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner  
 CC which is useful for the stimulation of a highly specific immune response  
 CC when administered to an animal. The protein of the invention may be  
 CC post-translationally modified (e.g. by the addition of fatty acids to  
 CC enhance immunogenicity). The fusion proteins of the invention can be  
 CC used as vaccines to induce an immune response. When a T cell epitope is  
 CC attached, they can be used for control of viral infections, bacterial  
 CC infections, parasitic infection and cancer. The fusion proteins can be  
 CC used in pharmaceutical compositions for the treatment of gastrointestinal  
 CC diseases, pulmonary infections, respiratory infections, and HIV  
 CC infections. The use of ubiquitin as a scaffold is also useful for the  
 CC generation of anti-gonadotropin releasing hormone antibodies which result  
 CC in the suppression of luteinizing hormone and follicle stimulating  
 CC hormone. This indirectly suppresses steriodogenesis and gamete maturation  
 CC in males and females. This type of anti-self response in humans is useful  
 CC in the treatment of prostate cancer and breast cancer. In livestock, the  
 CC ability to stimulate an anti-self response provides a simple alternative  
 CC to physical castration. Immunocastration of pigs is a better alternative  
 CC to physical castration, as it does not result in any of the detrimental  
 CC side effects associated with physical castration. Other examples of  
 CC diseases and conditions treated with self proteins fused with ubiquitin  
 CC are TNF and its epitopes to modulate septic shock, arthritis,  
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig  
 CC epsilon heavy chain for the control of allergic reactions; chorionic  
 CC gonadotropin for fertility control; and sperm proteins for fertility  
 CC control. A further use of the fusion proteins is as part of a vaccine to  
 CC enhance growth rate and thereby the final weight of the livestock prior  
 CC to shipment to market. In addition, the fusion proteins of the invention  
 CC can be used to detect and identify antibodies from experimental samples.  
 CC This sequence represents a GnRH fragment used in the construction of  
 CC a ubiquitin fusion protein described in the method of the invention.  
 CC  
 XX Sequence 40 AA:  
 SQ  
 Query Match 68.8%; Score 199.5; DB 20; Length 40;  
 Best Local Similarity 77.6%; Pred. No. 1e-15; 2; Indels 9; Gaps 3;  
 Matches 38; Conservative 0; Mismatches 2;  
 QY 1 OHWSYGLRPGSGQDWSYGLRPGSSQHMWSYGLRPGSGQDWSYGLRPG 49  
 Db 1 qhwsyglrpg---qhwsyglrpg---qhwsyglrpg---qhwsyglrpg 40  
 RESULT 14  
 ID Y31182 standard; peptide; 41 AA.  
 XX Y31182:  
 AC 28-OCT-1999 (first entry)  
 DT  
 DE Ubiquitin fusion protein GnRH fragment.  
 XX  
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;  
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;  
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;  
 KW pulmonary infection; gamete maturation; prostate; anti-self; pig;  
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;  
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;

KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;  
 XX fertility; sperm protein; growth rate; antibody; detection; GnRH.  
 OS Unidentified.  
 XX  
 XX WO9942472-A1.  
 PN  
 XX  
 PD 26-AUG-1999.  
 XX  
 XX 26-JAN-1999; 99WO-US01588.  
 PF  
 XX 19-FEB-1998; 98US-0026276.  
 PR  
 XX (IGEN-) IGEN INT INC.  
 PA  
 XX  
 PI Kenten JH, Lohmas GL, Pilon AL, Roberts SF, Tramontano A;  
 DR WPI; 1999-518582/43.  
 XX  
 XX  
 PT Epitope-containing fusion proteins used to generate a highly  
 XX specific immune responses  
 PS Claim 81; Page 43; 67pp; English.  
 XX  
 CC This invention describes a novel fusion protein, comprising a heat shock  
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner  
 CC which is useful for the stimulation of a highly specific immune response  
 CC when administered to an animal. The protein of the invention may be  
 CC post-translationally modified (e.g. by the addition of fatty acids to  
 CC enhance immunogenicity). The fusion proteins of the invention can be  
 CC used as vaccines to induce an immune response. When a T cell epitope is  
 CC attached, they can be used for control of viral infections, bacterial  
 CC infections, parasitic infection and cancer. The fusion proteins can be  
 CC used in pharmaceutical compositions for the treatment of gastrointestinal  
 CC diseases, pulmonary infections, respiratory infections, and HIV  
 CC infections. The use of ubiquitin as a scaffold is also useful for the  
 CC presentation and stimulation of anti-self immune responses, e.g.  
 CC generation of anti-gonadotropin releasing hormone antibodies which result  
 CC in the suppression of luteinizing hormone and follicle stimulating  
 CC hormone. This indirectly suppresses steroidogenesis and gamete maturation  
 CC in males and females. This type of anti-self response in humans is useful  
 CC in the treatment of prostate cancer and breast cancer. In livestock, the  
 CC ability to stimulate an anti-self response provides a simple alternative  
 CC to physical castration. Immunocastration of pigs is a better alternative  
 CC to physical castration, as it does not result in any of the detrimental  
 CC side effects associated with physical castration. Other examples of  
 CC diseases and conditions treated with self proteins fused with ubiquitin  
 CC are TNF and its epitopes to modulate septic shock, arthritis,  
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig  
 CC epsilon heavy chain for the control of allergic reactions; chorionic  
 CC gonadotropin for fertility control; and sperm proteins for fertility  
 CC control. A further use of the fusion proteins is as part of a vaccine to  
 CC enhance growth rate and thereby the final weight of the livestock prior  
 CC to shipment to market. In addition, the fusion proteins of the invention  
 CC can be used to detect and identify antibodies from experimental samples.  
 CC This sequence represents a GnRH fragment used in the construction of  
 CC a ubiquitin fusion protein described in the method of the invention.  
 XX  
 SO Sequence 41 AA;

Query Match 68.8%; Score 199.5; DB 20; Length 41;  
 Best Local Similarity 77.6%; Pred. No. 1,1e-15;  
 Matches 38; Conservative 0; Mismatches 2; Indels 9; Gaps 3;

OY 1 QHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49  
 Db 1 qhwsyglrpg---qhwsyglrpg---qhwsyglrpg---qhwsyglrpg 40

RESULT 15  
 ID R1187 standard; Protein; 323 AA.

XX  
 AC R1187;  
 XX  
 DT 22-MAR-1991 (first entry)  
 XX  
 DE Plasmid pBT8A859-encoded Tratp-multiple LHRH analogue fusion.  
 XX  
 KW Tratp protein; Luteinizing hormone releasing hormone; fusion protein;  
 XX immunological castration.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..20  
 FT /label= Tratp signal  
 FT 201..280  
 FT Peptide /label= 8 LHRH analogues in tandem repeat  
 XX  
 PN WO9102799-A.  
 XX  
 PD 07-MAR-1991.  
 XX  
 PF 24-AUG-1990; 90WO-AU00373.  
 XX  
 PR 25-AUG-1989; 89AU-0005979.  
 XX  
 PA (BIOF-) BIOTECHN AUST PTY L.  
 XX  
 PI Russell-Jones GJ, Stewart AG, Tsoulis CG;  
 XX  
 DR WPI; 1991-087282/12.  
 DR N-PSDB; Q11021.  
 XX  
 PT Fusion proteins comprising LHRH analogue and Tratp (analogue) -  
 PT useful in vaccine for inhibition or control of reproduction in  
 PT vertebrates, esp. domestic animals  
 XX  
 XX Example 1; Fig 2 and 5; 53pp; English.  
 PS  
 XX  
 CC Plasmid pBT8A859 is a Tratp-LHRH analogue fusion in which 8 copies  
 CC of an LHRH analogue have been inserted between amino acids 200 and  
 CC 201 of Tratp (Ogata R.T. et al., (1982) J.Bacteriol. 151:819-827).  
 CC The plasmid was constructed by two successive additions of DNA  
 CC coding for a dimer of LHRH analogue into the SmaI site of pBT8A862  
 CC (see Q11020) which all ready carries four copies of the LHRH  
 CC sequence. After transformation, colonies with 8 LHRH molecules were  
 CC identified. Fusion proteins with multiple inserts generated a higher  
 CC anti-LHRH response (as measured by the binding of (125)I-LHRH at a  
 CC serum dilution of 1:2000 final) than constructs with a single  
 CC insert, in outbred mice and dogs. The fusion proteins can be used to  
 CC inhibit reproductive functions in vertebrates.  
 CC See also Q10995, Q10997-Q11000, Q11014-Q11020.  
 XX  
 SO Sequence 323 AA;

Query Match 64.7%; Score 187.5; DB 12; Length 323;  
 Best Local Similarity 69.4%; Pred. No. 2e-13;  
 Matches 34; Conservative 4; Mismatches 2; Indels 9; Gaps 3;

OY 1 QHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49  
 Db 201 ehwsyglrpg---ehwsyglrpg---ehwsyglrpg---ehwsyglrpg 240

Search completed: March 2, 2001, 10:53:47  
 Job time: 912 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:59:03 ; Search time 201.99 Seconds

(without alignments)  
37.240 Million cell updates/sec

Title: US-09-306-689-11

Perfect score: 290  
Sequence: 1 QHWSYGLRPGSGSQDMSYGLRG.....SYGLRPGSGSQDMSYGLRPG 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 957798 seqs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/paa/PCNUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
27: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
28: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	100.0	49	US-09-019-010-4	Sequence 4, Appl1
2	290	100.0	49	US-09-249-447A-7	Sequence 7, Appl1
3	290	100.0	49	US-09-305-924-11	Sequence 11, Appl1
4	290	100.0	49	US-09-306-689-11	Sequence 11, Appl1
5	290	100.0	49	US-09-383-912-4	Sequence 4, Appl1

6	290	100.0	544	17	US-09-383-912-10	Sequence 10, Appl1
7	290	100.0	695	17	US-09-305-924-13	Sequence 13, Appl1
8	290	100.0	695	17	US-09-306-689-13	Sequence 13, Appl1
9	290	100.0	699	17	US-09-383-912-16	Sequence 16, Appl1
10	290	100.0	699	17	US-09-383-912-16	Sequence 16, Appl1
11	199.5	68.8	40	1	PCT-US00-22121-35	Sequence 8, Appl1
12	199.5	68.8	40	14	US-09-026-276-35	Sequence 35, Appl1
13	199.5	68.8	40	17	US-09-374-721A-35	Sequence 35, Appl1
14	199.5	68.8	41	1	PCT-US00-22121-34	Sequence 34, Appl1
15	199.5	68.8	41	14	US-09-026-276-34	Sequence 34, Appl1
16	199.5	68.8	41	17	US-09-374-721A-34	Sequence 34, Appl1
17	187.5	64.7	40	19	US-09-506-078-15	Sequence 15, Appl1
18	187.5	64.7	40	22	US-60-120-454-15	Sequence 15, Appl1
19	187.5	64.7	44	5	US-08-160-882-45	Sequence 45, Appl1
20	187.5	64.7	84	5	US-08-160-882-47	Sequence 47, Appl1
21	187.5	64.7	397	19	US-09-506-078-31	Sequence 31, Appl1
22	187.5	64.7	397	22	US-60-120-454-31	Sequence 31, Appl1
23	187.5	64.7	398	19	US-09-506-078-25	Sequence 25, Appl1
24	187.5	64.7	398	22	US-60-120-454-25	Sequence 25, Appl1
25	187.5	64.7	411	19	US-09-506-078-23	Sequence 23, Appl1
26	187.5	64.7	411	22	US-60-120-454-23	Sequence 23, Appl1
27	187.5	64.7	442	19	US-09-506-078-27	Sequence 27, Appl1
28	187.5	64.7	442	22	US-60-120-454-27	Sequence 27, Appl1
29	108	37.2	20	1	PCT-US00-22121-26	Sequence 26, Appl1
30	108	37.2	20	8	US-08-476-013-12	Sequence 12, Appl1
31	108	37.2	20	14	US-09-026-276-26	Sequence 26, Appl1
32	108	37.2	21	17	US-09-374-721A-26	Sequence 26, Appl1
33	108	37.2	21	8	US-08-477-298-14	Sequence 14, Appl1
34	106	36.6	20	3	US-07-761-849-10	Sequence 10, Appl1
35	106	36.6	20	8	US-08-476-013-10	Sequence 10, Appl1
36	105	36.2	20	1	PCT-US00-22121-30	Sequence 30, Appl1
37	105	36.2	20	1	PCT-US00-22121-31	Sequence 31, Appl1
38	105	36.2	20	14	US-09-026-276-30	Sequence 30, Appl1
39	105	36.2	20	14	US-09-026-276-31	Sequence 31, Appl1
40	105	36.2	20	17	US-09-374-721A-30	Sequence 30, Appl1
41	105	36.2	21	17	US-09-374-721A-31	Sequence 31, Appl1
42	105	36.2	21	3	US-07-761-849-11	Sequence 11, Appl1
43	105	36.2	21	3	US-07-761-849-12	Sequence 12, Appl1
44	105	36.2	21	8	US-08-476-013-11	Sequence 11, Appl1
45	105	36.2	21	8	US-08-476-013-12	Sequence 12, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-019-010-4  
: Sequence 4, Application US/09019010  
: GENERAL INFORMATION:  
: APPLICANT: HARLAND, RICHARD  
: APPLICANT: MANN, JOHN G.  
: APPLICANT: ACRES, STEPHEN D.  
: TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS  
: MOLECULES  
: NUMBER OF SEQUENCES: 6  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: ROBINS & ASSOCIATES  
: STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
: CITY: MENLO PARK  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94025  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/019,010  
: FILING DATE: 05-FEB-1998  
: CLASSIFICATION: 424  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 60/036,883

FILING DATE: 05-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0035  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 325-7812  
TELEFAX: (650) 325-7823  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-019-010-4

Query Match 100.0%; Score 290; DB 14; Length 49;  
Best Local Similarity 100.0%; Pred. No. 4.7e-24;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
Db 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49

RESULT 2  
US-09-249-447A-7  
Sequence 7, Application US/09249447A  
GENERAL INFORMATION:  
APPLICANT: Manns, Jack G.  
TITLE OF INVENTION: PASSIVE IMMUNIZATION AS A TREATMENT FOR HORMONE  
FILE REFERENCE: 9001-0045  
CURRENT APPLICATION NUMBER: US/09/249,447A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/075,637  
PRIOR FILING DATE: 1998-02-23  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: GnRH-2, Fig.  
US-09-249-447A-7

Query Match 100.0%; Score 290; DB 16; Length 49;  
Best Local Similarity 100.0%; Pred. No. 4.7e-24;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
Db 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49

RESULT 3  
US-09-305-924-11  
Sequence 11, Application US/09305924A  
GENERAL INFORMATION:  
APPLICANT: Jack G. Manns  
APPLICANT: Stephen D. Acres  
APPLICANT: Richard Harland  
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION  
FILE REFERENCE: 9001-0048  
CURRENT APPLICATION NUMBER: US/09/305,924A  
CURRENT FILING DATE: 1999-05-05  
EARLIER APPLICATION NUMBER: US 60/084,217  
EARLIER FILING DATE: 1998-05-05  
NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 49  
TYPE: PRT  
ORGANISM: GnRH  
US-09-305-924-11

Query Match 100.0%; Score 290; DB 17; Length 49;  
Best Local Similarity 100.0%; Pred. No. 4.7e-24;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
Db 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49

RESULT 4  
US-09-306-689-11  
Sequence 11, Application US/09306689B  
GENERAL INFORMATION:  
APPLICANT: Robbins, Sarah C.  
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN  
FILE REFERENCE: 9001-0047  
CURRENT APPLICATION NUMBER: US/09/306,689B  
CURRENT FILING DATE: 1999-05-06  
EARLIER APPLICATION NUMBER: US 60/088,024  
EARLIER FILING DATE: 1998-06-04  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Sequence  
US-09-306-689-11

Query Match 100.0%; Score 290; DB 17; Length 49;  
Best Local Similarity 100.0%; Pred. No. 4.7e-24;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
Db 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49

RESULT 5  
US-09-383-912-4  
Sequence 4, Application US/09383912  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANNs, JOHN G.  
TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/383,912  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-383-912-4

Query Match 100.0%; Score 290; DB 17; Length 49;  
Best Local Similarity 100.0%; Pred. No. 4.7e-24;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49  
|||||  
DB 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49

RESULT 6  
US-09-383-912-10  
Sequence 10, Application US/09383912  
GENERAL INFORMATION:  
APPLICANT: POTTER, ANDREW A.  
APPLICANT: MANN, JOHN G.  
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS LLP  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/383,912  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/694,865  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MCCracken, THOMAS P.  
REGISTRATION NUMBER: 38,548  
REFERENCE/DOCKET NUMBER: 9001-0016.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)327-3400  
TELEFAX: (415)327-3231  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 544 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-383-912-10

Query Match 100.0%; Score 290; DB 17; Length 544;

Best Local Similarity 100.0%; Pred. No. 5e-23;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49  
|||||  
DB 494 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 542

RESULT 7  
US-09-305-924-13  
Sequence 13, Application US/09305924A  
GENERAL INFORMATION:  
APPLICANT: Jack G. Manns  
APPLICANT: Stephen D. Acres  
APPLICANT: Richard Harland  
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION  
FILE REFERENCE: 9001-0048  
CURRENT APPLICATION NUMBER: US/09/305,924A  
CURRENT FILING DATE: 1999-05-05  
EARLIER APPLICATION NUMBER: US 60/084,217  
EARLIER FILING DATE: 1998-05-05  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 13  
LENGTH: 695  
TYPE: PRP  
ORGANISM: GNRH  
US-09-305-924-13

Query Match 100.0%; Score 290; DB 17; Length 695;  
Best Local Similarity 100.0%; Pred. No. 6.4e-23;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49  
|||||  
DB 9 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 57

RESULT 8  
US-09-306-689-13  
Sequence 13, Application US/09306689B  
GENERAL INFORMATION:  
APPLICANT: Robbins, Sarah C.  
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN  
FILE REFERENCE: 9001-0047  
CURRENT APPLICATION NUMBER: US/09/306,689B  
CURRENT FILING DATE: 1999-05-06  
EARLIER APPLICATION NUMBER: US 60/088,024  
EARLIER FILING DATE: 1998-06-04  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 13  
LENGTH: 695  
TYPE: PRP  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Sequence  
US-09-306-689-13

Query Match 100.0%; Score 290; DB 17; Length 695;  
Best Local Similarity 100.0%; Pred. No. 6.4e-23;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49  
|||||  
DB 9 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 57

RESULT 9

US-09-383-912-16  
; Sequence 16, Application US/09383912  
; GENERAL INFORMATION:  
; APPLICANT: POTTER, ANDREW A.  
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: REED & ROBINS LLP  
; STREET: 285 HAMILTON AVENUE, SUITE 200  
; CITY: PALO ALTO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/383,912  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/694,865  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MCCracken, THOMAS P.  
; REGISTRATION NUMBER: 38,548  
; REFERENCE/DOCKET NUMBER: 9001-0016.22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)327-3400  
; TELEFAX: (415)327-3231  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 699 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-383-912-16

Query Match 100.0%; Score 290; DB 17; Length 699;  
Best Local Similarity 100.0%; Pred. No. 6.4e-23;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
DB 9 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 57

RESULT 10  
US-09-383-912-8  
; Sequence 8, Application US/09383912  
; GENERAL INFORMATION:  
; APPLICANT: POTTER, ANDREW A.  
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: REED & ROBINS LLP  
; STREET: 285 HAMILTON AVENUE, SUITE 200  
; CITY: PALO ALTO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/383,912

US-09-383-912-8  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/694,865  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MCCracken, THOMAS P.  
; REGISTRATION NUMBER: 38,548  
; REFERENCE/DOCKET NUMBER: 9001-0016.22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)327-3400  
; TELEFAX: (415)327-3231  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 977 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-383-912-8

Query Match 100.0%; Score 290; DB 17; Length 977;  
Best Local Similarity 100.0%; Pred. No. 9e-23;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
DB 927 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 975

RESULT 11  
PCT-US00-22121-35  
; Sequence 35, Application PC/TUS0022121  
; GENERAL INFORMATION:  
; APPLICANT: Proteinix Company  
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM  
; FILE REFERENCE: IGN-2004WO  
; CURRENT APPLICATION NUMBER: PCT/US00/22121  
; CURRENT FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 09/026,276  
; PRIOR FILING DATE: 1998-02-19  
; PRIOR APPLICATION NUMBER: US 09/374,721  
; PRIOR FILING DATE: 1999-08-13  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 35  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide  
; PCT-US00-22121-35

Query Match 68.8%; Score 199.5; DB 1; Length 40;  
Best Local Similarity 77.6%; Pred. No. 1.6e-14;  
Matches 38; Conservative 0; Mismatches 2; Indels 9; Gaps 3;

QY 1 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49  
DB 1 QHWSYGLRPG---QHWSYGLRPG---QHWSYGLRPG---QHWSYGLRPG 40

RESULT 12  
US-09-026-276-35  
; Sequence 35, Application US/09026276  
; GENERAL INFORMATION:  
; APPLICANT: Kenten, John H  
; APPLICANT: Tramontano, Alfonso  
; APPLICANT: Pilon, Aprile L  
; APPLICANT: Lohmas, Gerald F  
; APPLICANT: Roberts, Steven F

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	68.8%	Score 199.5	DB 14	Length 40
Best Local Similarity	77.6%	Pred. No. 1.6e14		
Matches 38	Conservative 0	Mismatches 2	Indels 9	Gaps 3

```

OY      1 QHWSYGLRPGSGSQDWSTYGLRPGSSQHWSYGLRPGSGSQDWSTYGLRPG 49
          |||||  | |||||  |||||  | |||||
          1 QHWSYGLRPG--QHWSYGLRPG--QHWSYGLRPG 400
DB

```

```

RESULT 13
US-09-374-721A-35
: Sequence 35. Application US/09374721A
: GENERAL INFORMATION:
: APPLICANT: Kenten, John H.
: APPLICANT: Roberts, Steven
: APPLICANT: Lohmas, Gerald
: TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
: FILE REFERENCE: CIP OF ICN-9601
: CURRENT APPLICATION NUMBER: US/09/374,721A
: CURRENT FILING DATE: 1999-08-13
: PRIOR APPLICATION NUMBER: 09/026,276
: PRIOR FILING DATE: 1998-02-19
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 35
: LENGTH: 40
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: polypeptide
US-09-374-721A-35

```

Query Match	68.8%	DB 17	Length 40
Best Local Similarity	77.6%	Pred. No. 1.6e-14	
Matches 38; Conservative	0	Mismatches 2	Indels 9; Gaps 3

```

OY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSGSQDWSYGLRPG 49
    |||||  |||||  |||||  |||||
Db 1 QHWSYGLRPG--QHWSYGLRPG--QHWSYGLRPG--QHWSYGLRPG 40

```

```

RESULT 14
PCT-US00-22121-34
Sequence 34, Application PC/TUS0022121
GENERAL INFORMATION:
APPLICANT: Proteinix Company
TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
FILE REFERENCE: IGN-2004MO
CURRENT APPLICATION NUMBER: PCT/US00/22121
PRIORITY FILING DATE: 2000-08-14
PRIORITY FILING DATE: 09/026,.276
PRIORITY FILING DATE: 1998-02-19
PRIORITY APPLICATION NUMBER: US 09/374,721
PRIORITY FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 34
LENGTH: 41

```

```

: TYPE: PRT
:
: ORGANISM: Artificial Sequence
:
: FEATURE:
:
: OTHER INFORMATION: Description of Artificial Sequence: polypeptide
:
: OTHER INFORMATION: antigen
:
PCT-US00-22121-34

```

Query Match	68.8%	Score 199.5	DB 1	Length 41
Best Local Similarity	77.6%	Pred. No. 1.6e-14		
Matches 38; Conservative	0	Mismatches 2	Indels 9	Gaps 3

```
QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSGSQDWSYGLRPG 493
      |||||  | |||||  |||||  | |||||
Db 1 QHWSYGLRPG---QHWSYGLRPG---QHWSYGLRPG 400
```

```

RESULT 15
US-09-026-276-34
: Sequence 34, Application US/09026276
: GENERAL INFORMATION:
: APPLICANT: kentee, John H
: APPLICANT: Tramontano, Alfonso
: APPLICANT: Pilon, April L
: APPLICANT: Lohnas, Gerald F
: APPLICANT: Roberts, Steven F
: TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED SYSTEM
: FILE REFERENCE: U.S. Patent Application No. 09/026,276
: CURRENT APPLICATION NUMBER: US/09/026,276
: CURRENT FILING DATE: 1998-02-19
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 34
: LENGTH: 41
: TYPE: PRT
: ORGANISM: Porcine
: US-09-026-276-34

```

Query Match	68.8%	Score 199.5	DB 14	Length 41
Best Local Similarity	77.6%	Pred. No. 1.6e-14		
Matches 38	0	Mismatches 2	Indels 9	Gaps 3

```
QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSGSQDWSYGLRPG 49
      |||||  |||||  |||||  |||||
Db 1 QHWSYGLRPG---QHWSYGLRPG---QHWSYGLRPG 400
      |||||  |||||  |||||  |||||
```

Search completed: March 2, 2001, 10:59:04  
Job time: 374 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 2, 2001, 10:44:35 ; Search time 47.48 Seconds  
(without alignments)  
5.295 Million cell updates/sec

Title: US-09-306-689-2

Perfect score: 58

Sequence: 1 XHMSYXLRPGXXXX 14

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/PCtUS.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/Backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	87.9	10	1	US-07-714-540-9
2	51	87.9	10	1	US-07-690-983D-2
3	51	87.9	10	1	US-07-690-983D-32
4	51	87.9	10	1	US-08-103-022-1
5	51	87.9	10	1	US-08-127-351-6
6	51	87.9	10	1	US-08-184-935-6
7	51	87.9	10	1	US-08-193-374-3
8	51	87.9	10	1	US-08-193-374-8
9	51	87.9	10	1	US-08-343-883-1
10	51	87.9	10	1	US-08-480-367B-6
11	51	87.9	10	1	US-08-000-931-5
12	51	87.9	10	1	US-08-487-221A-6
13	51	87.9	10	1	US-08-480-370-6
14	51	87.9	10	1	US-08-428-468-22
15	51	87.9	10	1	US-08-341-219-11
16	51	87.9	10	1	US-08-453-588-2
17	51	87.9	10	1	US-08-453-588-4
18	51	87.9	10	1	US-08-453-588-6
19	51	87.9	10	1	US-08-453-588-8
20	51	87.9	10	1	US-08-453-588-10
21	51	87.9	10	1	US-08-453-588-12
22	51	87.9	10	1	US-08-453-588-14
23	51	87.9	10	1	US-08-453-588-16
24	51	87.9	10	1	US-08-453-588-19
25	51	87.9	10	1	US-08-453-588-22
26	51	87.9	10	1	US-08-188-223-3
27	51	87.9	10	1	US-08-406-935-5
28	51	87.9	10	1	US-08-591-917-1

29	51	87.9	10	1	US-08-387-156-2	Sequence 2, Appli
30	51	87.9	10	1	US-08-474-555-1	Sequence 1, Appli
31	51	87.9	10	1	US-08-446-692-1	Sequence 1, Appli
32	51	87.9	10	1	US-08-242-678D-1	Sequence 1, Appli
33	51	87.9	10	1	US-08-242-678D-7	Sequence 7, Appli
34	51	87.9	10	2	US-08-796-598-6	Sequence 6, Appli
35	51	87.9	10	2	US-08-694-865-2	Sequence 2, Appli
36	51	87.9	10	2	US-08-694-865-18	Sequence 18, Appli
37	51	87.9	10	2	US-08-488-351A-1	Sequence 1, Appli
38	51	87.9	10	2	US-08-480-494B-1	Sequence 1, Appli
39	51	87.9	10	2	US-08-447-175A-6	Sequence 6, Appli
40	51	87.9	10	3	US-08-878-748-2	Sequence 2, Appli
41	51	87.9	10	3	US-08-521-079-2	Sequence 2, Appli
42	51	87.9	10	3	US-08-521-079-4	Sequence 4, Appli
43	51	87.9	10	3	US-08-521-079-6	Sequence 6, Appli
44	51	87.9	10	3	US-08-521-079-8	Sequence 8, Appli
45	51	87.9	10	5	5492893-1	Patent No. 5492893

## ALIGNMENTS

RESULT 1  
US-07-714-540-9  
Sequence 9, Application US/07714540  
Patent No. 5262521  
GENERAL INFORMATION:  
APPLICANT: Almqvist, Ronald G.  
TITLE OF INVENTION: ISOLATED ATRIAL PEPTIDE-DEGRADING  
TITLE OF INVENTION: ENZYME AND NOVEL COMPOUNDS USEFUL AS INHIBITORS THEREOF  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESS: Irell & Manella  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: California  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07714,540  
FILING DATE: 19910607  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Reed, Dianne E.  
REGISTRATION NUMBER: 31,292  
REFERENCE/DOCKET NUMBER: 8500-0135.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-327-7250  
TELEFAX: 415-327-2951  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-714-540-9

Query Match 87.9% Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 HMSYXLRPG 10  
|||||  
DB 2 HMSYXLRPG 10

RESULT 2  
US-07-690-983D-2  
Sequence 2, Application US/07690983D  
Patent No. 5403586  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J.  
APPLICANT: STEWART, Andrew G.  
APPLICANT: TSONIS, Con G.  
TITLE OF INVENTION: FUSION PROTEINS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/690,983D  
CLASSIFICATION: 435  
FILING DATE: 25-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU90/00373  
FILING DATE: 24-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/148 CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-07-690-983D-2

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 3  
US-07-690-983D-32  
Sequence 32, Application US/07690983D  
Patent No. 5403586  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, Gregory J.  
APPLICANT: STEWART, Andrew G.  
APPLICANT: TSONIS, Con G.  
TITLE OF INVENTION: FUSION PROTEINS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/690,983D  
FILING DATE: 25-JUN-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU90/00373  
FILING DATE: 24-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/148 CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-690-983D-32

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 4  
US-08-103-022-1  
Sequence 1, Application US/08103022  
Patent No. 5413990  
GENERAL INFORMATION:  
APPLICANT: Haviv, Fortuna  
APPLICANT: Fitzpatrick, Timothy D.  
APPLICANT: Swenson, Rolf E.  
APPLICANT: Nichols, Charles J.  
APPLICANT: Mort, Nicholas A.  
TITLE OF INVENTION: N-Terminus Modified Analogs of LHRH  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377  
STREET: Abbott Laboratories, One Abbott Park Road  
CITY: No. 5413990th Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/103,022  
FILING DATE: 05-OCT-1993  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Janssen, Jerry F.  
REGISTRATION NUMBER: 29,175  
REFERENCE/DOCKET NUMBER: 5389.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708) 938-7742  
TELEFAX: (708) 938-2623  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids



TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /note="Xaa at position 1 is a  
OTHER INFORMATION: 5-oxo-prolyl aminoacyl residue."  
US-08-103-022-1

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 5  
US-08-127-351-6  
Sequence 6, Application US/08127351  
Patent No. 5449761  
GENERAL INFORMATION:  
APPLICANT: BELINKA JR, BENJAMIN A.  
APPLICANT: COUGHLIN, DANIEL J.  
APPLICANT: ALVAREZ, VERNON L.  
APPLICANT: WOOD, RICHARD  
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE  
TITLE OF INVENTION: CONSTRUCTS  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
ADDRESSEE: NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127/351  
FILING DATE: 28-SEP-1993  
CLASSIFICATION: 534  
ATTORNEY/AGENT INFORMATION:  
NAME: Villacorta, Gilberto M.  
REGISTRATION NUMBER: 34,038  
REFERENCE/DOCKET NUMBER: 4980-004-44  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 6  
OTHER INFORMATION: /note="Position 6 may be either  
OTHER INFORMATION: Gly or D-Trp."  
US-08-127-351-6

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 2 HWSYXLRPG 10

RESULT 6  
US-08-184-935-6  
Sequence 6, Application US/08184935  
Patent No. 5476770  
GENERAL INFORMATION:  
APPLICANT: PRADELLES, PHILIPPE  
TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN  
TITLE OF INVENTION: OR HAPTEN  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/184,935  
FILING DATE: 24-JAN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5476770man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 846-286-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /note="C-terminal amide"  
US-08-184-935-6

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 7  
US-08-193-374-3  
Sequence 3, Application US/08193374  
Patent No. 5487898  
GENERAL INFORMATION:  
APPLICANT: Fu Lu, Mou-Ying  
APPLICANT: Reiland, Thomas L.  
TITLE OF INVENTION: Compositions and Method for the

TITLE OF INVENTION: Sublingual or Buccal Administration of Therapeutic Agents  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377 - AP6D  
STREET: Abbott Laboratories, One Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,374  
FILING DATE: 07-FEB-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07 / 983,111  
FILING DATE: 30-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Janssen, Jerry F.  
REGISTRATION NUMBER: 29,175  
REFERENCE/DOCKET NUMBER: 4848.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708) 938-7742  
TELEFAX: (708) 938-2623  
INFORMATION FOR SEQ. ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /note="XAA at position 1 is a  
OTHER INFORMATION: 5-oxopropyl residue."  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 6  
OTHER INFORMATION: /note="XAA at position 6 is a  
OTHER INFORMATION: D-tryptyl residue."  
US-08-193-374-3

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
DB 2 HWSYXLRPG 10

RESULT 8  
US-08-193-374-8  
Sequence 8, Application US/08193374  
Patent No. 5487898  
GENERAL INFORMATION:  
APPLICANT: Fu Lu, Mou-Ying  
TITLE OF INVENTION: Compositions and Method for the  
TITLE OF INVENTION: Sublingual or Buccal Administration of Therapeutic Agents  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377 - AP6D  
STREET: Abbott Laboratories, One Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA

ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,374  
FILING DATE: 07-FEB-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07 / 983,111  
FILING DATE: 30-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Janssen, Jerry F.  
REGISTRATION NUMBER: 29,175  
REFERENCE/DOCKET NUMBER: 4848.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708) 938-7742  
TELEFAX: (708) 938-2623  
INFORMATION FOR SEQ. ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /note="XAA at position 1 is a  
OTHER INFORMATION: 5-oxopropyl residue."  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 6  
OTHER INFORMATION: /note="XAA at position 6 is a  
OTHER INFORMATION: D-3-(naphth-2-yl)alanyl residue."  
US-08-193-374-8

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
DB 2 HWSYXLRPG 10

RESULT 9  
US-08-343-883-1  
Sequence 1, Application US/08343883  
Patent No. 5573767  
GENERAL INFORMATION:  
APPLICANT: Dufour, Raymond J.  
APPLICANT: Roulet, Claude J.M.  
APPLICANT: Chouvet, Claire D.  
APPLICANT: Bonneau, Michel B.  
TITLE OF INVENTION: Method for improving the organoleptic  
TITLE OF INVENTION: qualities of the meat from uncastated male domestic  
TITLE OF INVENTION: animals, vaccines which are usable in this method, new  
TITLE OF INVENTION: peptide, in particular for producing these vaccines...  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Larson and Taylor  
STREET: 727 Twenty-Third Street, South  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/343,883  
FILING DATE: 17-NOV-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/946,495  
FILING DATE: 09-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 9102513  
FILING DATE: 01-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 9115289  
FILING DATE: 10-DEC-1991  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 10  
OTHER INFORMATION: /label= NH2  
OTHER INFORMATION: /note= "amidated glycine"  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1  
OTHER INFORMATION: /label= pyro  
OTHER INFORMATION: /note= "pyroglutamic acid"  
PUBLICATION INFORMATION:  
AUTHORS: Matsuo, H.  
AUTHORS: Baba, Y.  
AUTHORS: G. Nair, R. M.  
AUTHORS: Arimura, A.  
AUTHORS: Schally, A. V.  
TITLE: Structure of the porcine LH- and  
TITLE: FSH-releasing hormone. I. The proposed amino acid  
TITLE: Sequence.  
JOURNAL: Biochem. Biophys. Res. Commun.  
VOLUME: 43  
ISSUE: 6  
PAGES: 1334-1339  
DATE: 1971  
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 10  
US-08-343-883-1

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 HWSYXLRPG 10  
||| |||  
Db 2 HWSYGLRPG 10

RESULT 10  
US-08-480-367B-6  
Sequence 6, Application US/08480367B  
Patent No. 5578288  
GENERAL INFORMATION:  
APPLICANT: BELINKA JR, BENJAMIN A.  
APPLICANT: COUGHLIN, DANIEL J.  
APPLICANT: ALVAREZ, VERNON L.  
APPLICANT: WOOD, RICHARD  
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE  
TITLE OF INVENTION: CONSTRUCTS  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER  
STREET: 99 Canal Center Plaza, Suite 300

CITY: Alexandria  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,367B  
FILING DATE: 07-06-95  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Villacorta, Gilberto M.  
REGISTRATION NUMBER: 34,038  
REFERENCE/DOCKET NUMBER: 2654-002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 684-1111  
TELEFAX: (703) 684-1124  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 6  
OTHER INFORMATION: /note= "Position 6 may be either  
OTHER INFORMATION: Gly or D-Tip."  
US-08-480-367B-6

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYXLRPG 10  
||| |||  
Db 2 HWSYXLRPG 10

RESULT 11  
US-08-000-931-5  
Sequence 5, Application US/08000931  
Patent No. 5578477  
GENERAL INFORMATION:  
APPLICANT: TAMANOI DR., FUYUHIKO  
TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF  
TITLE OF INVENTION: INHIBITORS OF PROTEIN FARNESYLTRANSFERASE  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/000,931  
FILING DATE: 05-JAN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 64098/102/ARDE  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-000-931-5

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
DB 2 HWSYXLRPG 10

RESULT 12  
US-08-487-221A-6  
Sequence 6, Application US/08487221A  
Patent No. 5593656  
GENERAL INFORMATION:  
APPLICANT: BELINKA JR, BENJAMIN A.  
APPLICANT: COUGHLIN, DANIEL J.  
APPLICANT: ALVAREZ, VERNON L.  
APPLICANT: WOOD, RICHARD  
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE  
TITLE OF INVENTION: CONSTRUCTS  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &  
ADDRESSEE: NEUSTADT, P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,221A  
CLASSIFICATION: 424  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,351  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Villacorta, Gilberto M.  
REGISTRATION NUMBER: 34,038  
REFERENCE/DOCKET NUMBER: 4980-004-44  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 6  
OTHER INFORMATION: /note= "Position 6 may be either  
OTHER INFORMATION: Gly or D-Trp."

US-08-487-221A-6

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
DB 2 HWSYXLRPG 10

RESULT 13  
US-08-480-370-6  
Sequence 6, Application US/08480370  
Patent No. 560847  
GENERAL INFORMATION:  
APPLICANT: BELINKA JR, BENJAMIN A.  
APPLICANT: COUGHLIN, DANIEL J.  
APPLICANT: ALVAREZ, VERNON L.  
APPLICANT: WOOD, RICHARD  
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE  
TITLE OF INVENTION: CONSTRUCTS  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &  
ADDRESSEE: NEUSTADT, P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,370  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,351  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Villacorta, Gilberto M.  
REGISTRATION NUMBER: 34,038  
REFERENCE/DOCKET NUMBER: 4980-004-44  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 6  
OTHER INFORMATION: /note= "Position 6 may be either  
OTHER INFORMATION: Gly or D-Trp."

US-08-480-370-6

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
DB 2 HWSYXLRPG 10

Db 2 HWSYXLRPG 10

RESULT 14  
US-08-428-488-22  
Sequence 22, Application US/08428488  
Patent No. 5624894  
GENERAL INFORMATION:  
APPLICANT: BODOR, Nicholas S.  
TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE  
PEPTIDES BY SEQUENTIAL METABOLISM  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P. O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428.488  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Mary Katherine  
REGISTRATION NUMBER: 26,254  
REFERENCE/DOCKET NUMBER: 028724-087  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /note= "Position 1 = p-Glu."  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /note= "Position 10 = Gly-NH2."  
US-08-428-488-22

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
Db 2 HWSYXLRPG 10

RESULT 15  
US-08-341-219-11  
Sequence 11, Application US/08341219  
Patent No. 5643877  
GENERAL INFORMATION:  
APPLICANT: Zohar, Y.  
APPLICANT: Rivier, J.  
APPLICANT: Powell, J.  
APPLICANT: Sherwood, N.  
APPLICANT: Gothliff, Y.  
TITLE OF INVENTION: Compounds and Methods For Controlling

TITLE OF INVENTION: Reproduction in Fish  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: N.Y.  
COUNTRY: USA  
ZIP: 10036-2211  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/341.219  
FILING DATE: 05-DEC-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30742  
REFERENCE/DOCKET NUMBER: 8399-003-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /product= "OTHER"  
OTHER INFORMATION: /label= Glu1  
OTHER INFORMATION: /note= "pyroglutamic acid"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /product= "OTHER"  
OTHER INFORMATION: /label= Gly10  
OTHER INFORMATION: /note= "amidated"  
US-08-341-219-11

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.00095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
Db 2 HWSYXLRPG 10

Search completed: March 2, 2001, 10:54:36  
Job time: 601 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 2, 2001, 10:52:50 ; Search time 201.99 Seconds  
(without alignments)  
10.640 Million cell updates/sec

Title: US-09-306-689-2

Perfect score: 58  
Sequence: 1 XHMSYXLRPCXXXX 14

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 957798 seqs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Pending-Patents-AA:\*

- 1: /cgn2\_6/ptodata/2/paa/PCRTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/PCRNEM\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 27: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 28: /cgn2\_6/ptodata/2/paa/US50\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	89.7	9	20	US-09-657-276-160
2	52	89.7	10	20	US-09-657-276-139
3	51	87.9	9	20	US-09-657-276-153
4	51	87.9	10	1	PCT-US00-22121-28
5	51	87.9	10	1	PCT-US00-22121-32

6	51	87.9	10	1	PCT-US94-04832A-1	Sequence 1, Appl
7	51	87.9	10	1	PCT-US94-12763-4	Sequence 4, Appl
8	51	87.9	10	1	PCT-US94-13394-5	Sequence 5, Appl
9	51	87.9	10	1	PCT-US95-01225-3	Sequence 3, Appl
10	51	87.9	10	1	PCT-US96-16950-2	Sequence 2, Appl
11	51	87.9	10	1	PCT-US96-16950-3	Sequence 3, Appl
12	51	87.9	10	1	PCT-US96-16950-7	Sequence 7, Appl
13	51	87.9	10	1	PCT-US96-16950-8	Sequence 8, Appl
14	51	87.9	10	1	PCT-US96-16950-9	Sequence 9, Appl
15	51	87.9	10	1	PCT-US96-17008-3	Sequence 3, Appl
16	51	87.9	10	1	PCT-US96-17008-7	Sequence 7, Appl
17	51	87.9	10	1	PCT-US96-17008-8	Sequence 8, Appl
18	51	87.9	10	1	PCT-US96-17008-9	Sequence 9, Appl
19	51	87.9	10	1	PCT-US96-17008-10	Sequence 10, Appl
20	51	87.9	10	1	PCT-US96-17008-13	Sequence 13, Appl
21	51	87.9	10	1	PCT-US99-11219-1143	Sequence 1143, Ap
22	51	87.9	10	1	PCT-US99-11219-1309	Sequence 1309, Ap
23	51	87.9	10	1	PCT-US99-11219-1344	Sequence 1344, Ap
24	51	87.9	10	1	PCT-US99-13960-28	Sequence 28, Appl
25	51	87.9	10	1	PCT-US99-13975B-77	Sequence 77, Appl
26	51	87.9	10	3	US-07-669-695-1	Sequence 1, Appl
27	51	87.9	10	3	US-07-672-300A-14	Sequence 14, Appl
28	51	87.9	10	3	US-07-728-782A-1	Sequence 1, Appl
29	51	87.9	10	3	US-07-728-782-1	Sequence 1, Appl
30	51	87.9	10	3	US-07-761-849-1	Sequence 1, Appl
31	51	87.9	10	3	US-07-946-062-22	Sequence 22, Appl
32	51	87.9	10	3	US-07-984-293-5	Sequence 5, Appl
33	51	87.9	10	4	US-08-020-366-9	Sequence 9, Appl
34	51	87.9	10	4	US-08-020-985-9	Sequence 9, Appl
35	51	87.9	10	4	US-08-057-166-1	Sequence 1, Appl
36	51	87.9	10	5	US-08-138-514-4	Sequence 4, Appl
37	51	87.9	10	5	US-08-138-514-11	Sequence 11, Appl
38	51	87.9	10	5	US-08-138-514-12	Sequence 12, Appl
39	51	87.9	10	5	US-08-138-514-16	Sequence 16, Appl
40	51	87.9	10	5	US-08-138-514-20	Sequence 20, Appl
41	51	87.9	10	5	US-08-138-516-1	Sequence 1, Appl
42	51	87.9	10	5	US-08-138-516-6	Sequence 6, Appl
43	51	87.9	10	5	US-08-138-516-7	Sequence 7, Appl
44	51	87.9	10	5	US-08-138-516-11	Sequence 11, Appl
45	51	87.9	10	27	US-09-412-558-1	Sequence 1, Appl

## ALIGNMENTS

US-09-657-276-160  
: Sequence 160, Application us/09657276  
: GENERAL INFORMATION:  
: APPLICANT: Conjuchem, Inc.  
: APPLICANT: Bridon, Dominique  
: APPLICANT: Ezrin, Alan  
: APPLICANT: Milner, Peter  
: APPLICANT: Holmes, Darren  
: APPLICANT: Thibaudau, Karen  
: TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM  
: TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD  
: TITLE OF INVENTION: COMPONENTS  
: FILE REFERENCE: 2110  
: CURRENT APPLICATION NUMBER: US/09/657,276  
: CURRENT FILING DATE: 2000-09-07  
: PRIOR APPLICATION NUMBER: 60/134,406  
: PRIOR FILING DATE: 1999-05-17  
: PRIOR APPLICATION NUMBER: 60/153,406  
: PRIOR FILING DATE: 1999-09-10  
: PRIOR APPLICATION NUMBER: 60/159,783  
: PRIOR FILING DATE: 1999-10-18  
: NUMBER OF SEQ ID NOS: 1617  
: SOFTWARE: Patent Ver. 2.1  
: SEQ ID NO 160  
: LENGTH: 9  
: TYPE: PRT  
: ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptide  
US-09-657-276-160

Query Match 89.7%; Score 52; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 8.3e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 1 HWSYGLRPG 9

RESULT 2  
US-09-657-276-139  
Sequence 139, Application US/09657276

GENERAL INFORMATION:  
APPLICANT: Conjuchem, Inc.  
APPLICANT: Bridon, Dominique  
APPLICANT: Ezrin, Alan  
APPLICANT: Milner, Peter  
APPLICANT: Holmes, Darren  
TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM  
TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD  
FILE REFERENCE: 2110

CURRENT APPLICATION NUMBER: US/05/657, 276  
CURRENT FILING DATE: 2000-09-07  
PRIOR APPLICATION NUMBER: 60/134, 406  
PRIOR FILING DATE: 1999-05-17  
PRIOR APPLICATION NUMBER: 60/153, 406  
PRIOR FILING DATE: 1999-09-10  
PRIOR APPLICATION NUMBER: 60/155, 783  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 1617  
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 139  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-657-276-139

Query Match 89.7%; Score 52; DB 20; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.015;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 2 HWSYALRPG 10

RESULT 3  
US-09-657-276-153  
Sequence 153, Application US/09657276

GENERAL INFORMATION:  
APPLICANT: Conjuchem, Inc.  
APPLICANT: Bridon, Dominique  
APPLICANT: Ezrin, Alan  
APPLICANT: Milner, Peter  
APPLICANT: Holmes, Darren  
TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM  
TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD  
FILE REFERENCE: 2110  
CURRENT APPLICATION NUMBER: US/09/657, 276

CURRENT FILING DATE: 2000-09-07  
PRIOR APPLICATION NUMBER: 60/134, 406  
PRIOR FILING DATE: 1999-05-17  
PRIOR APPLICATION NUMBER: 60/153, 406  
PRIOR FILING DATE: 1999-09-10  
PRIOR APPLICATION NUMBER: 60/159, 783  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 1617  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 153

LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-657-276-153

Query Match 87.9%; Score 51; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 8.3e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 1 HWSYGLRPG 9

RESULT 4  
PCT-US00-22121-28  
Sequence 28, Application PC/TUS0022121

GENERAL INFORMATION:  
APPLICANT: Proteinix Company  
TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM  
FILE REFERENCE: IGN-2004MO  
CURRENT APPLICATION NUMBER: PCT/US00/22121  
CURRENT FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 09/026, 276  
PRIOR FILING DATE: 1998-02-19  
PRIOR APPLICATION NUMBER: US 09/374, 721  
PRIOR FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 28  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: polypeptide  
PCT-US00-22121-28

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
Db 2 HWSYGLRPG 10

RESULT 5  
PCT-US00-22121-32  
Sequence 32, Application PC/TUS0022121

GENERAL INFORMATION:  
APPLICANT: Proteinix Company  
TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM  
FILE REFERENCE: IGN-2004MO  
CURRENT APPLICATION NUMBER: PCT/US00/22121  
CURRENT FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 09/026, 276  
PRIOR FILING DATE: 1998-02-19



PRIOR APPLICATION NUMBER: US 09/374,721  
PRIOR FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: Patentl Ver. 2.0  
SEQ ID NO 32  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: polypeptide  
PCT-US00-22121-32

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 2 HWSYGLRPG 10

RESULT 6  
PCT-US94-04832A-1  
Sequence 1, Application PC/TUS9404832A  
GENERAL INFORMATION:

APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang YI  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulatores for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE:  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: US  
ZIP: 11530

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentl Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/04832A  
FILING DATE: 13-APR-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME:  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516)742-4343  
TELEFAX: (516)742-4366  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-04832A-1

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 2 HWSYGLRPG 10

RESULT 7  
PCT-US94-12763-4  
Sequence 4, Application PC/TUS9412763  
GENERAL INFORMATION:

APPLICANT:  
TITLE OF INVENTION: Nucleic Acid Encoding [His-5,Tyr-7,Tyr-8]-  
TITLE OF INVENTION: GnRH Preprohormone and [Ser-6]-GnRH preprohormone and  
TITLE OF INVENTION: Their Uses  
NUMBER OF SEQUENCES: 31  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentl Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/12763  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/147,771  
FILING DATE: 05-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 14210-000400PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Mammal

FEATURE: Peptide  
NAME/KEY: 1..10  
LOCATION: 1..10  
OTHER INFORMATION: /note="GnRH"  
PCT-US94-12763-4

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
DB 2 HWSYGLRPG 10

RESULT 8  
PCT-US94-13394-5  
Sequence 5, Application PC/TUS9413394  
GENERAL INFORMATION:

APPLICANT: Seang H. Yiv  
TITLE OF INVENTION: Transparent Liquid for  
TITLE OF INVENTION: Encapsulating Drug Delivery  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and  
ADDRESS: Norris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/13394  
FILING DATE: herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 885,202  
FILING DATE: May 20, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: David R. Bailey  
REGISTRATION NUMBER: 35,057  
REFERENCE/DOCKET NUMBER: AFB1-0349  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acid residues  
TYPE: Amino Acid  
STRANDEDNESS:  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: pyroglutamic acid  
PCT-US94-13394-5

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||||  
Db 2 HWSYGLRPG 10

RESULT 9  
PCT-US95-01225-3  
Sequence 3, Application PC/TUS9501225  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Scibienski, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Divas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01225  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Divas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label= pglu  
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..10  
OTHER INFORMATION: /note= "immunomimic"  
PCT-US95-01225-3

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||||  
Db 2 HWSYGLRPG 10

RESULT 10  
PCT-US96-16950-2  
Sequence 2, Application PC/TUS9616950  
GENERAL INFORMATION:  
APPLICANT: LOMBARDO, VICTORIA K.  
APPLICANT: MARBURG, STEPHEN  
APPLICANT: TOLMAN, RICHARD L.  
TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING  
TITLE OF INVENTION: HORMONE  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MOLLIE M. YANG  
STREET: 126 E. LINCOLN AVE PO BOX 2000  
CITY: RAHWAY  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/16950  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: YANG, MOLLIE M.  
REGISTRATION NUMBER: 32,718  
REFERENCE/DOCKET NUMBER: 19444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-6343  
TELEFAX: 908-594-4720  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
PCT-US96-16950-2

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||||  
DB 2 HWSYXLRPG 10

## RESULT 11

PCT-US96-16950-3

Sequence 3, Application PC/TUS9616950

GENERAL INFORMATION:  
APPLICANT: LOMBARDO, VICTORIA K.  
APPLICANT: MARBURG, STEPHEN  
APPLICANT: TOLMAN, RICHARD L.  
TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MOLLIE M. YANG  
STREET: 126 E. LINCOLN AVE PO BOX 2000  
CITY: RAHWAY  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/16950  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: YANG, MOLLIE M.  
REGISTRATION NUMBER: 32,718  
REFERENCE/DOCKET NUMBER: 19444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-6343  
TELEFAX: 908-594-4720  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
PCT-US96-16950-3

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||||  
DB 2 HWSYXLRPG 10

## RESULT 12

PCT-US96-16950-7

Sequence 7, Application PC/TUS9616950

GENERAL INFORMATION:  
APPLICANT: LOMBARDO, VICTORIA K.  
APPLICANT: MARBURG, STEPHEN  
APPLICANT: TOLMAN, RICHARD L.  
TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MOLLIE M. YANG  
STREET: 126 E. LINCOLN AVE PO BOX 2000  
CITY: RAHWAY

STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/16950  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: YANG, MOLLIE M.  
REGISTRATION NUMBER: 32,718  
REFERENCE/DOCKET NUMBER: 19444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-6343  
TELEFAX: 908-594-4720  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
PCT-US96-16950-7

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10  
|||||  
DB 2 HWSYXLRPG 10

## RESULT 13

PCT-US96-16950-8

Sequence 8, Application PC/TUS9616950

GENERAL INFORMATION:  
APPLICANT: LOMBARDO, VICTORIA K.  
APPLICANT: MARBURG, STEPHEN  
APPLICANT: TOLMAN, RICHARD L.  
TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MOLLIE M. YANG  
STREET: 126 E. LINCOLN AVE PO BOX 2000  
CITY: RAHWAY  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/16950  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: YANG, MOLLIE M.  
REGISTRATION NUMBER: 32,718  
REFERENCE/DOCKET NUMBER: 19444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-6343  
TELEFAX: 908-594-4720  
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
PCT-US96-16950-8

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
DB 2 HWSYXLRPG 10

RESULT 14  
PCT-US96-16950-9  
Sequence 9, Application PC/TUS9616950  
GENERAL INFORMATION:  
APPLICANT: LOMBARDO, VICTORIA K.  
APPLICANT: MARBURG, STEPHEN  
APPLICANT: TOLMAN, RICHARD L.  
TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELASING  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MOLLIE M. YANG  
STREET: 126 E. LINCOLN AVE PO BOX 2000  
CITY: RAHWAY  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/16950  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: YANG, MOLLIE M.  
REGISTRATION NUMBER: 32,718  
REFERENCE/DOCKET NUMBER: 19444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-6343  
TELEFAX: 908-594-4720  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
PCT-US96-16950-9

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
DB 2 HWSYXLRPG 10

RESULT 15

PCT-US96-17008-3  
Sequence 3, Application PC/TUS9617008  
GENERAL INFORMATION:  
APPLICANT: Hickey, Gerard J.  
APPLICANT: Mohr, Kenneth L.  
TITLE OF INVENTION: PSEUDOMONAS EXOTOXIN AS IMMUNOGENIC  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MOLLIE M. YANG  
STREET: 126 E. Lincoln Ave. P.O. Box 2000  
CITY: RAHWAY  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/17008  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Yang, Mollie M.  
REGISTRATION NUMBER: 32,718  
REFERENCE/DOCKET NUMBER: 19445  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-6343  
TELEFAX: 908-594-4720  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US96-17008-3

Query Match 87.9%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10  
DB 2 HWSYXLRPG 10

Search completed: March 2, 2001, 10:59:02  
Job time: 372 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:45 ; Search time 66.16 Seconds  
(without alignments)  
8.786 Million cell updates/sec

Title: US-09-306-689-3

Perfect score: 108

Sequence: 1 CPPPSEHMSYGLRPG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.GeneSeq\_36:\*

1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	17	16	R78282
2	108	100.0	17	21	F58139
3	88.5	81.9	16	16	R78284
4	81	75.0	34	21	Y91244
5	81	75.0	34	21	Y68594
6	81	75.0	35	21	Y91242
7	81	75.0	35	21	Y91243
8	81	75.0	35	21	Y68593
9	71	65.7	49	17	W03944
10	71	65.7	49	19	W79567
11	71	65.7	49	19	W61542
12	71	65.7	49	21	Y58363

13	71	65.7	49	21	Y58135
14	71	65.7	544	17	W03943
15	71	65.7	544	19	W79570
16	71	65.7	695	19	W79573
17	71	65.7	695	21	Y58361
18	71	65.7	695	21	Y58133
19	71	65.7	977	17	W03942
20	71	65.7	977	19	W79569
21	69	63.9	92	7	P61427
22	68	63.0	253	12	R11181
23	68	63.0	263	12	R11185
24	68	63.0	323	12	R11187
25	66	61.1	28	15	R62698
26	66	61.1	256	12	R11177
27	65	60.2	29	16	R74262
28	65	60.2	253	12	R11184
29	64	59.3	11	19	W69276
30	64	59.3	16	15	R44956
31	64	59.3	936	14	R34547
32	63	58.3	10	2	P10411
33	63	58.3	10	2	P10416
34	63	58.3	10	6	P50222
35	63	58.3	10	7	P60127
36	63	58.3	10	7	P61403
37	63	58.3	10	7	P60576
38	63	58.3	10	8	P70922
39	63	58.3	10	10	P90630
40	63	58.3	10	12	R15713
41	63	58.3	10	13	R26819
42	63	58.3	10	15	R62689
43	63	58.3	10	16	R91197
44	63	58.3	10	16	R68645
45	63	58.3	10	21	Y53061

## ALIGNMENTS

RESULT 1	
R78282	R78282 standard; peptide; 17 AA.
XX	
AC	R78282;
XX	
XX	13-MAR-1996 (first entry)
DE	GnRH immunomimetic and spacer (GnRH(1-10)-Ser1).
XX	
KW	Immunomimetic; gonadotropin releasing hormone; GnRH; LHRH; vaccine;
KW	luteinizing hormone releasing hormone; GnRH; LHRH; vaccine;
KW	gonadal steroid hormone associated dependent disease; gonadotropin;
KW	immunological contraception; mammal; breast; cancer; prostate;
KW	benign prostatic hypertrophy.
XX	
OS	Chimeric - Homo sapiens.
OS	Chimeric - Synthetic.
XX	
PH	Key
FT	Peptide
FT	Peptide
FT	Peptide
FT	Modified-site
XX	
PN	W09520600-AI.
PD	03-AUG-1995.
XX	
XX	26-JAN-1995; 95WO-US01225.
PF	GnRH-2. Synthetic
XX	Peptide hormone Gn
PR	Four-copy gonadotr
XX	
XX	27-JAN-1994; 94US-0188223.

GnRH analogue mult  
LKT-GnRH protein f  
LKT-GnRH chimeric  
LKT-GnRH chimeric  
Leukotoxin/gonadot  
Gonadotropin relea  
LKT-GnRH protein f  
LKT-GnRH chimeric  
Human preproLHRH P  
Plasmid pBR733-en  
Plasmid pBR733-en  
Plasmid pBR733-en  
LHRH-containing im  
Plasmid pBR733-en  
SSAL2 TH2LHRH. Sy  
Plasmid pBR733-en  
Luteinizing hormo  
E. coli fimbriae H  
GnRH-leukotoxin ge  
Luteinizing Hormon  
Luteinizing Hormon  
Gonadotropin rele  
Gonadoliberin anta  
Gonadotropin relea  
Novel decapeptide  
Luteinizing Hormon  
Sequence of lutein  
Peptide #1 with ho  
LH releasing hormo  
LHRH happen for at  
LHRH peptide. Syn  
Gonadotropin relea  
Luteinizing hormon

XX (APHT-) APHTON CORP.  
XX Grimes S, Scibienski R;  
XX WPI; 1995-275410/36.  
XX New anti-gonadotropin releasing hormone immunogenic composition(s)  
PT - used for treating gonadotropin and gonadal steroid hormone  
PT dependent disease(s) and providing contraception.  
XX Claim 1; Page 29; 39pp; English.  
XX The sequences given in R78282-85 represent immunomimetics to  
CC gonadotropin releasing hormone (GnRH). These peptides comprise the  
CC wild type GnRH sequence and a spacer attached to either the N- or C-  
CC terminal. The spacer molecules serve as a link through which the  
CC immunomimetic is attached to an immunological carrier such as  
CC diphtheria toxoid (DT) and also affects the immune response generated  
CC by the vaccinated mammal against the immunomimetic. Compositions  
CC comprising these peptides may be used for treating a mammal for  
CC gonadotropin and gonadal steroid hormone associated dependent  
CC disease or for providing immunological contraception in mammals.  
CC They can also be used for treating breast cancer, uterine and other  
CC gynaecological cancers, endometriosis, uterine fibroids, prostate  
CC cancer, or benign prostatic hypertrophy.  
SQ Sequence 17 AA;

Query Match 100.0%; Score 108; DB 16; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CPPPSSSEHWSYGLRPG 17  
Db 1 cpppssehsyglrpg 17  
|||||

RESULT 2  
Y58139  
ID Y58139 standard; peptide: 17 AA.  
XX Y58139;  
XX 07-MAR-2000 (first entry)  
XX Gonadotropin releasing hormone (GnRH) peptide analogue 3.  
XX Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;  
KW antibody; immunogenic; chimeric; testosterone; androgenic;  
KW non-androgenic; steroid; reduction; weight gain; muscle distribution;  
KW fat distribution; male pattern; boar taint; flavour; impairment;  
KW reliable; immunocastration; meat production.  
OS Synthetic.  
XX WO956771-AZ.  
XX 11-NOV-1999.  
XX 05-MAY-1999; 99WO-CA00360.  
XX 05-MAY-1998; 98US-0084217.  
XX (BIOS-) BIOSSTAR INC.  
XX Manns JG, Acres SD, Harland R;  
XX WPI; 2000-062125/05.  
XX Production of uncastrated male food animals using vaccines -  
XX PT

PS Disclosure; Page 11; 87pp; English.  
XX Sequences Y58136-Y58141 represent gonadotropin releasing hormone  
CC (GnRH) analogues which may be used as an alternative to sequence  
CC Y58135 in embodiments of the present invention. The invention  
CC relates to a method of using two GnRH immunogen vaccines to produce  
CC uncastrated male animals for meat production, one vaccination prior to  
CC or during the fattening period to reduce circulating testosterone levels,  
CC and the second vaccination about 2-8 weeks before slaughter to  
CC substantially reduce androgenic and/or non-androgenic steroids. The  
CC invention is used to produce food animals that exhibit the weight gain  
CC and muscle/fat distribution of male animals without the problems  
CC associated with male animals. Such problems include "boar taint", a  
CC urine-like odour found in cooked meat of uncastrated pigs which is  
CC caused by steroids stored in the tissues, and similar flavour  
CC impairments in the meat of other intact male animals. The invention is  
CC more reliable than prior art immunocastration techniques.  
SQ Sequence 17 AA;

Query Match 100.0%; Score 108; DB 21; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CPPPSSSEHWSYGLRPG 17  
Db 1 cpppssehsyglrpg 17  
|||||

RESULT 3  
R78284  
ID R78284 standard; peptide: 16 AA.  
XX R78284;  
XX 13-MAR-1996 (first entry)  
XX GnRH immunomimetic and spacer (GnRH(1-10)-Arg1).  
XX Immunomimetic; gonadotropin releasing hormone; GnRH; LHRH; vaccine;  
KW luteinising hormone releasing hormone; spacer; immunomimetic; uterine;  
KW diphtheria toxoid; DT; gynaecological; endometriosis; uterine fibroids;  
KW gonadal steroid hormone associated dependent disease; gonadotropin;  
KW immunological contraception; mammal; breast; cancer; prostate;  
KW benign prostatic hypertrophy.  
XX Chimeric - Homo sapiens.  
OS Chimeric - Synthetic.  
XX Key Location/Qualifiers  
FH Peptide 1..6  
FT /note= "spacer"  
FT Peptide 7..16  
FT /note= "GnRH"  
FT Modified-site 16  
FT /note= "Amidated C-terminal"  
XX WO9520600-A1.  
XX 03-AUG-1995.  
XX 26-JAN-1995; 95WO-US01225.  
XX 27-JAN-1994; 94US-0188223.  
XX (APHT-) APHTON CORP.  
XX Grimes S, Scibienski R;  
XX WPI; 1995-275410/36.  
XX New anti-gonadotropin releasing hormone immunogenic composition(s)  
XX PT

PT - used for treating gonadotropin and gonadal steroid hormone  
PR dependent disease(s) and providing contraception.  
XX  
PS Claim 1: Page 29; 39pp; English.  
XX  
CC The sequences given in R78282-85 represent immunomodulators to  
CC gonadotropin releasing hormone (GnRH). These peptides comprise the  
CC wild type GnRH sequence and a spacer attached to either the N- or C-  
CC terminal. The spacer molecules serve as a link through which the  
CC immunomodulator is attached to an immunological carrier such as  
CC diphtheria toxin (DT) and also affects the immune response generated  
CC by the vaccinated mammal against the immunomodulator. Compositions  
CC comprising these peptides may be used for treating a mammal for  
CC gonadotropin and gonadal steroid hormone associated dependent  
CC disease or for providing immunological contraception in mammals.  
CC They can also be used for treating breast cancer, uterine and other  
CC gynaecological cancers, endometriosis, uterine fibroids, prostate  
CC cancer, or benign prostatic hypertrophy.  
XX  
SQ Sequence 16 AA:  
  
Query Match 81.9%; Score 88.5; DB 16; Length 16;  
Best Local Similarity 88.2%; Pred. No. 2.1e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
  
OY 1 CPPPSSEHWSYGLRPG 17  
| | | | | | | | | | | | | | | | | |  
DB 1 CPPPP-rehwsyglrpg 16  
  
RESULT 4  
Y91244  
ID Y91244 standard; peptide; 34 AA.  
XX  
AC Y91244;  
XX  
DT 22-MAY-2000 (first entry)  
XX  
DE Modified HBV surface Ag/LHRH antigenic peptide, SEQ ID NO:122.  
XX  
XX Promiscuous T-cell epitope; measles virus F protein; MVF;  
KM hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
KM interleukin B virus releasing hormone; LHRH; contraceptive; anticancer;  
KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
KM foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;  
KM Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
KM cholesterol ester transport protein; anti-arteriosclerotic.  
XX  
OS Chimeric - Hepatitis B virus.  
XX  
PN WO966957-A2.  
XX  
PD 29-DEC-1999.  
XX  
PF 21-JUN-1999; 99WO-US13975.  
XX  
PR 20-JUN-1998; 98US-0100412.  
XX  
PA (UNBI-) UNITED BIOMEDICAL INC.  
XX  
PI Wang CY;  
XX  
DR WPI; 2000-160564/14.  
XX  
XX New artificial T helper cell epitope and derived immunogens with target  
PR antigenic site, for immunization against e.g. malaria, arteriosclerosis  
PT or human immune deficiency virus  
XX  
PS Disclosure: Page 109; 129pp; English.  
XX  
CC The invention relates to novel promiscuous T helper cell epitopes (Th),  
CC and immunogenic peptides comprising the Th epitopes of the invention

CC along with B cell epitopes. The Th epitopes and peptide immunogens  
CC containing them, are used to induce a T helper cell response.  
CC specifically against Plasmodium falciparum, cholesterol ester transport  
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,  
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
CC peptide immunogens may be used for prevention and/or treatment of  
CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
CC immunotherapy; for inhibition of the action of interleukin hormone  
CC releasing hormone (LHRH) for contraception, treatment of hormone-  
CC dependent cancer, prevention of boar taint in meat, and  
CC immunocastration); for promoting the growth of animals; or for  
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous  
CC Th (functional in genetically diverse subjects) into an immunogen  
CC improves capacity to induce a strong T helper cell-mediated immune  
CC response, resulting in production of antibodies against a target antigen.  
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.  
CC Sequence Y91121 represents a promiscuous T helper epitope from the  
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and  
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th  
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from  
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are  
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,  
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH  
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target  
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is  
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising  
CC somatostatin and a Th epitope. Somatostatin immunogens may be used  
CC to promote growth in livestock. Y91208 is a human CD4 CDR-1-like domain  
CC antigenic site, and Y91209-Y90211 are MVA Th epitope/CD4 CDR2  
CC antigenic peptides which may be used to prevent HIV infection of T  
CC cells. Y90212 is a modified version of a human IGE (immunoglobulin  
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic  
CC peptides which may be used in the treatment of allergies. Y91220 is  
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid  
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223  
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and  
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be  
CC used in a malaria vaccine. Y91228-Y91231 represent CERP-derived peptides  
CC and Y91232-Y91241 are immunogens comprising a CERP peptide and a Th  
CC epitope which may be used to prevent or treat arteriosclerosis and  
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising  
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic  
CC peptides comprising MVA Th and HIV-1 B-cell epitope which may be used as  
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively  
CC an immunostimulatory intrasin protein epitope from Yerestina species, and  
CC hinge spacer peptide, both of which may optionally be used in the  
CC antigenic peptides of the invention.  
XX  
SQ Sequence 34 AA:  
  
Query Match 75.0%; Score 81; DB 21; Length 34;  
Best Local Similarity 81.2%; Pred. No. 0.00038;  
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
OY 2 PPPPSSEHWSYGLRPG 17  
| | | | | | | | | | | | | | | | | |  
DB 19 ppxpxpshwsyglrpg 34  
  
RESULT 5  
Y68594  
ID Y68594 standard; peptide; 34 AA.  
XX  
AC Y68594;  
XX  
DT 05-MAY-2000 (first entry)  
XX  
DE Peptide immunogen comprising a Th epitope and LHRH target antigen.  
XX  
KM Helper T cell epitope; surface antigen; peptide immunogen; LHRH;  
XX interleukin hormone-releasing hormone; spermatogenesis; ovulation;  
KM oestrus; sexual development; sex hormone; promiscuous T helper epitope;

Query Match	75.0%;	Score 81;	DB 21;	Length 34;
Best Local Similarity	81.2%;	Pred. No. 0.00038;		
Matches 13;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps	0;			

  

Key	Location/Qualifiers
Peptide	1..18
/note=	"helper Th epitope"
Peptide	19..24
/note=	"spacer Y68564"
Misc-difference	21
/note=	"not specified"
Misc-difference	23
/note=	"not specified"
Peptide	25..34
/note=	"LHRH antigenic epitope Y68566"

  

W09966952-A1.
29-DEC-1999.
21-JUN-1999;
99WO-US13960.
20-JUN-1998;
98US-0100414.
(UNBI-) UNITED BIOMEDICAL INC.
Wang CY;
WPI: 2000-160562/14.

  

New peptide immunogen containing luteinising hormone-releasing hormone antigen site and helper T cell epitope, for e.g. contraception and treatment of cancer

Example 1; Page 85; 102pp; English.

The present sequence represents a peptide immunogen comprising a synthetic helper T cell (Th) epitope and a target antigen, luteinising hormone-releasing hormone (LHRH). The synthetic Th epitope is derived from a structured synthetic antigen library (SSAL) designated SSAL2 Th2. SSAL2 Th2 is modeled after a promiscuous epitope taken from the Hepatitis B virus surface antigen. The peptide immunogens cause induction of a specific immune response to LHRH which is involved in regulation of spermatogenesis, ovulation, oestrus, sexual development and secretion of sex hormones. Provision of a promiscuous T helper epitope (which is functional in genetically diverse subjects) provides optimum immunogenicity to the B cell epitopes of the target antigen and thus high antibody titres against the target antigen. The peptide immunogens of the invention are used to vaccinate against mammalian LHRH, for use as (reversible) contraceptive; control of hormone-dependent tumours (cancer of prostate or breast, also endometriosis); to prevent boar taint (and improve meat quality) and for immunocastration.



50 Sequence 35 AA;

```

QY      2 PPPPSSEHWSYGLRPG 17
          |||  |||||
Db      20 ppxpxpehwsyglrpg 35

```

RESULT	7
Y91243	
ID	Y91243 standard; peptide; 35 AA

DT 22-MAY-2000 (first entry)

DE	Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:121.
DE	

KM Promionosis T-cell epitope; measles virus F protein; MVE;  
KM hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
KM interleukin hormone releasing hormone; LHRH; contraceptive; anticancer  
KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
KM foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;  
KM plasmidum falciptarum; ciliocystozoicite; antimalarial; CEMP;  
KM cholesterol ester transport protein; anti-arteriosclerotic.

05 Chimeric - Measles virus.

PN W09966957-A2.

PD 29-DEC-1999

PF 21-JUN-1999; 99WO-US13975:  
YY

PR 20-JUN-1998; 98US-0100412  
VV

PA (UNBI-) UNITED BIOMEDICAL INC.  
XX

PI Wang CY  
YY

DR WPT; 2000-160564/14.  
YY

PT New artificial T helper cell epitope and derived immunogens with target  
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis  
PT or human immune deficiency virus -

PS Disclosure; Page 108; 129pp; English.

CC The invention relates to novel promiscuous T helper cell epitopes (Th),  
CC and immunogenic peptides comprising the Th epitopes of the invention  
CC along with B cell epitopes. The Th epitopes and peptide immunogens  
CC containing them, are used to induce a T helper cell response,  
CC conferring 'memory' to the response.

CC specifically against Plasmodium falciparum, cholesteryl ester transport  
CC protein (CETP) or HIV epitopes, but more generally against any pathogen  
CC immunoreactive self-antigen or tumour antigen. The 7b epitopes and  
CC peptide immunogens may be used for prevention and/or treatment of  
CC infections (HIV, foot-and-mouth disease or malaria); for cancer

immunotherapy; for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration); for promoting the growth of animals; or for treating allergies or arteriosclerosis, incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence Y91121 represents a promiscuous T helper epitope from the measles virus F (MF) protein and sequences Y91122-Y91142, Y91226 and Y91245-Y91246 represent synthetic Th epitopes based on the MF Th epitope. Sequence Y91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are synthetic epitopes derived from this HBV epitope. Y91156-Y91196, Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. Y91200 is somatostatin, and Y91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. Y91208 is a human CD4 CD82-like domain antigenic site, and Y91209-Y90211 are MVA Th epitope/CD4 CD82 antigenic peptides which may be used to prevent HIV infection of T cells. Y90212 is a modified version of a human IGE (immunoglobulin E) CH3 domain, and Y90213-Y90219 are Th epitope/IGE CH3 antigenic peptides which may be used in the treatment of allergies. Y91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and Y91224-Y91225 comprise the CS antigen and an MVA Th epitope and may be used in a malaria vaccine. Y91228-Y91231 represent CEMP-derived peptides and Y91232-Y91241 are immunogens comprising a CEMP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising B-cell epitopes, and Y91248-Y91251 and Y91256-Y91273 are antigenic peptides comprising MVA Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively an immunostimulatory invasion protein epitope from *Yersinia* species, and a hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

**SQ Sequence 35 AA;**

Query Match	75.0%	Score 81	DB 21	Length 35
Best Local Similarity	81.2%	Pred. No. 0.00039		
Matches 13; Conservative		0; Mismatches 3;	Indels 0;	Gaps 0;

```
QY      2 PPPSSEHWSYGLRPG 17
          ||| | | | | | | |
Db      20 ppxpxpehwsyglrpg 35
```

RESULT	8
Y68593	
ID	Y68593 standard; peptide; 35 AA.

AC Y68593;

DT 05-MAY-2000 (first entry)

DE Peptide immunogen comprising a Th epitope and LHRH target antigen

KM oestrus; sexual development; sex hormone, promiscuous T helper epitope; KM vaccine; contraceptive; hormone-dependent tumour; prostate cancer; KM breast cancer; endometritis; boar taint; meat quality; immunocastration

05	Synthetic.
05	Measles virus
05	Unidentified.
xx	

```

FH Key Location/Qualifiers
FT Peptide 1..19
FT /note= "helper Th epitope Y68549"
FT Misc-difference 4
FT /label= Ser, Thr
FT Misc-difference 7
FT /label= Lys, Arg
FT Misc-difference 8
FT /label= Gly, Thr
FT Misc-difference 12
FT /label= His, Thr
FT Misc-difference 13
FT /label= Lys, Arg
FT Misc-difference 16
FT /label= Gly, Thr
FT Peptide 20..25
FT /note= "spacer Y68564"
FT Misc-difference 22
FT /note= "not specified"
FT Misc-difference 24
FT /note= "not specified"
FT Peptide 26..35
FT /note= "LHRH antigenic epitope Y68566"
XX
XX WO966952-A1.
XX
XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13960.
XX
XX 20-JUN-1998; 98US-0100414.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY;
XX
XX WPI; 2000-160562/14.
XX
XX
XX New peptide immunogen containing luteinising hormone-releasing hormone
XX antigen site and helper T cell epitope, for e.g. contraception and
XX treatment of cancer -
XX
XX Example 1; Page 85; 102pp; English.
XX
XX The present sequence represents a peptide immunogen comprising a
XX synthetic helper T cell (Th) epitope and a target antigen, luteinising
XX hormone-releasing hormone (LHRH). The synthetic Th epitope is derived
XX from a structured synthetic antigen library (SSAL) designated SSAL1 Th1.
XX SSAL Th1 is modelled after a promiscuous epitope taken from the F protein
XX of the Measles virus. The peptide immunogens cause induction of a
XX specific immune response to LHRH which is involved in regulation of a
XX spermatogenesis, ovulation, oestrus, sexual development and secretion
XX of sex hormones. Provision of a promiscuous T helper epitope (which is
XX functional in genetically diverse subjects) provides optimum
XX immunogenicity to the B cell epitopes of the target antigen and thus
XX high antibody titres against the target antigen. The peptide immunogens
XX of the invention are used to vaccinate against mammalian LHRH, for use
XX as (reversible) contraceptive; control of hormone-dependent tumours
XX (cancer of prostate or breast, also endometriosis); to prevent boar
XX taint (and improve meat quality) and for immunocastration.
XX
XX Sequence 35 AA;

```

```

Query Match 75.0%; Score 81; DB 21; Length 35;
Best Local Similarity 81.2%; Pred. No. 0.00039;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 PPSSEHMSYGLRPG 17
DB 20 ppxpxpnhwsyglrpg 35

```

```

RESULT 9
ID W03944
XX W03944 standard; Protein; 49 AA.
AC W03944;
XX
XX 20-NOV-1996 (first entry)
XX
XX GnRH 4-repeat sequence.
XX
XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
XX fusion protein; immunogen; vaccine; fertility control;
XX KW
XX KW contraception; sterilisation; PCB113; PCB11.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Peptide 1..10
FT /label= GnRH
FT Peptide 11..13
FT /label= Spacer
FT Peptide 14..23
FT /label= GnRH
FT Peptide 24..26
FT /label= Spacer
FT Peptide 27..36
FT /label= GnRH
FT Peptide 37..39
FT /label= Spacer
FT Peptide 40..49
FT /label= GnRH
XX
XX WO9624675-A1.
XX
XX 15-AUG-1996.
XX
XX 24-JAN-1996; 96WO-CA00049.
XX
XX 10-FEB-1995; 95US-0387156.
XX
XX (UYSA-) UNIV SASKATCHEWAN.
XX
XX Manns JG, Potter AA;
XX
XX WPI; 1996-384447/38.
XX
XX N-PSDB; T37178.
XX
XX Gonadotropin-releasing hormone multimer fusion proteins - with
XX PT leukotoxin polypeptide for increased immunogenicity, useful in
XX PT antifertility vaccine prodn.
XX
XX Example 2; Fig 1B; 87pp; English.
XX
XX A synthetic DNA sequence (T37178) codes for a gonadotropin
XX releasing hormone (GnRH) tetramer (W03944), in which the 4 GnRH
XX repeat units are separated by spacers designed to increase
XX immunogenicity. The DNA sequence was incorporated into vector
XX pAA352 (ATCC 68283), which contains a truncated leukotoxin
XX gene (LKT 352) derived from Pasteurella hemolytica, to
XX give plasmid PCB113 (T37176). Escherichia coli transformants
XX CC produce an LKT-GnRH fusion protein (see also W03942 and W03943)
XX CC useful as a vaccine for fertility control of domestic or farm animals.
XX
XX Sequence 49 AA;

```

```

Query Match 65.7%; Score 71; DB 17; Length 49;
Best Local Similarity 80.0%; Pred. No. 0.0099;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 PPSSEHMSYGLRPG 17
DB 22 pgsqssghmsyglrpg 36

```

```

RESULT 10
W79567
ID W79567 standard; Protein: 49 AA.
XX
AC W79567;
XX
DT 24-DEC-1998 (first entry)
XX
DE GnRH-2.
XX
KM Gonadotropin releasing hormone; GnRH; decapeptide; hypothalamus; spacer;
KM LH: luteinizing hormone; FSH; follicle stimulating hormone; vertebrate;
KM pyroglu; chimera; leukotoxin polypeptide; multimer; vaccine; tumour;
KM immunogenic.
XX
OS Synthetic.
XX
PN MO9806848-A1.
XX
PD 19-FEB-1998.
XX
PF 08-AUG-1997; 97WO-CA00559.
XX
PR 09-AUG-1996; 96US-0694865.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Manns JG, Potter AA;
XX
DR WPI: 1998-159540/14.
XX
N-PSDB: V61529.
XX
PT Chimeric protein of leukotoxin and gonadotropin releasing hormone -
PT useful for, e.g. preparation of vaccines for reduction of incidence
PT of mammary tumours in mammals
XX
PS Disclosure: Figure 1B; 118pp; English.
XX
XX
CC The present sequence represents a recombinantly produced or chemically
CC synthesised Gonadotropin releasing hormone-2 (GnRH-2) polypeptide, which
CC contains four copies of the GnRH decapeptide (V61528) and triplet amino
CC acid spacers between each of these sequences. This decapeptide is
CC secreted naturally by the hypothalamus which controls release of both the
CC luteinizing hormone (LH) and the follicle stimulating hormone (FSH) in
CC vertebrates. This sequence, as compared to the native peptide, has been
CC found to have an N-terminal Gln rather than a pyroglu residue, and also
CC contains substitutions at amino acid residues 15 and 41, whereby His is
CC replaced by Asp. This produces an alternating multimeric GnRH sequence
CC which is highly immunogenic that can be used in the construction of a
CC chimeric protein that comprises a leukotoxin polypeptide, several
CC multimers, and the GnRH sequence. The chimeric protein can be used as a
CC vaccine to help reduce the incidence of mammary tumours in a mammalian
CC individual.
XX
SQ Sequence 49 AA:

```

```

Query Match 65.7%; Score 71; DB 19; Length 49;
Best Local Similarity 80.0%; Pred. No. 0.0099;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 PPPSSRHSYGLRPG 17
   | 11:|||||11111
DB 22 pygssqhwsgylrpg 36

```

```

RESULT 11
W61542
ID W61542 standard; Protein: 49 AA.
XX
AC W61542;

```

```

XX
DT 27-OCT-1998 (first entry)
XX
DE Peptide hormone GnRH-2 decapeptide (4 copies) fragment.
XX
KM GnRH: gonadotropin releasing hormone peptide hormone; leukotoxin;
KM immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;
KM immune response; hormone receptor; cancerous cell; domestic animal;
KM porcine; bovine; luteinizing hormone; follicle stimulating hormone;
KM immunocastrate.
XX
OS Synthetic.
XX
PN WO9834639-A1.
XX
PD 13-AUG-1998.
XX
PF 04-FEB-1998; 98WO-CA00059.
XX
PR 05-FEB-1997; 97US-0036883.
XX
PA (BIOS-) BIOSTAR INC.
XX
PI Acres SD, Harland R, Manns JG;
XX
DR WPI: 1998-446952/38.
XX
N-PSDB: V45190.
XX
PT Immunisation against endogenous molecules by administering vaccine
PT to ear - useful to elicit efficient and uniform immune response
PT against e.g. gonadotropin releasing hormone to immunocastrate pigs
PT and cattle
XX
PS Example 1; Fig 1B: 61pp; English.
XX
XX
CC This represents the amino acid sequence of the gonadotropin releasing
CC hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-
CC GnRH polypeptide gene fusions. This is used to exemplify the method of
CC invention of immunisation against endogenous molecules by administering
CC a vaccine which comprises an immunogen and a carrier to the ear of the
CC mammal. The method is useful for eliciting an efficient and uniform
CC immune response to block or suppress the activity of an endogenous
CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,
CC or to elicit an immune response against a targeted endogenous cell type
CC (e.g. a cancerous or otherwise diseased cell). It is especially useful
CC to reduce the levels of GnRH in domestic animals, especially in porcine
CC or bovine species. The use of GnRH immunogens in the vaccine reduces the
CC levels of luteinizing hormone and follicle stimulating hormone and helps
CC in immunocastrating the animal. Administration of vaccine compositions to
CC the ear instead of intramuscular administration into the neck increases
CC the efficiency of vaccination of mammals against endogenous immunogens,
CC and may increase uniformity of vaccine presentation since the ear is
CC relatively uniform from animal to animal.
XX
SQ Sequence 49 AA:

```

```

Query Match 65.7%; Score 71; DB 19; Length 49;
Best Local Similarity 80.0%; Pred. No. 0.0099;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 PPPSSRHSYGLRPG 17
   | 11:|||||11111
DB 22 pygssqhwsgylrpg 36

```

```

RESULT 12
Y58363
ID Y58363 standard; Protein: 49 AA.
XX
AC Y58363;
XX
DT 27-MAR-2000 (first entry)

```

XX DE Four-copy gonadotropin-releasing hormone (GnRH) multimer.  
 XX XX GnRH multimer; gonadotropin-releasing hormone; immunosterilisation;  
 KW immunoncontraception; vaccine; feline; canine; equine; cervine; ds  
 XX OS Mammalia.  
 OS Synthetic.  
 XX MO9962545-A2.  
 XX XX  
 XX PD 09-DEC-1999.  
 XX PF 28-MAY-1999; 99WO-CA00493.  
 XX PR 04-JUN-1998; 98US-0088024.  
 XX PR 06-MAY-1999; 99US-0306689.  
 XX PA (BIOS-) BIOSTAR INC.  
 XX PI Robbins SC;  
 XX PI WPI; 2000-086857/07.  
 XX DR N-PSDB; 255702.  
 XX DR  
 XX PT Hormone immunogens, analogues or antibodies used to manufacture  
 PT vaccines for suppression of reproductive behavior and fertility in  
 PT vertebrates -  
 XX PS Claim 5; Fig 5B; 88pp; English.  
 XX CC This sequence represents a four-copy gonadotropin-releasing  
 CC hormone (GnRH) multimer, where the second and fourth GnRH sequence  
 CC have a His to Asp substitution at position 2 of the GnRH sequence.  
 CC The invention relates to GnRH immunogens, analogues or antibodies  
 CC that cross-react with endogenous GnRH of a vertebrate. A specifically  
 CC claimed immunogenic fusion protein (Y58361) comprises, in the N to  
 CC C-terminal direction, a synthetic peptide sequence (Y58364), an eight  
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer  
 CC sequence of Y58363), the LKT protein (which functions as a carrier  
 CC protein), and a second eight copy GnRH multimer. The fusion protein may  
 CC be used in a vaccine composition for prepubertal administration to a  
 CC vertebrate subject to result in prolonged suppression of reproductive  
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are  
 CC used to manufacture a composition or vaccine for immunosterilisation or  
 CC immunoncontraception of feline, canine, equine or cervine subjects.  
 CC The vaccines are used to suppress reproductive behaviour and/or  
 CC fertility for at least 10 months. The prepubertal administration  
 CC results in a prolonged, long-term suppression of testicular development  
 CC and/or function in males, or a prolonged, long-term suppression of  
 CC ovarian development and/or function in females. The methods provide a  
 CC viable and desirable alternative to surgical forms of sterilisation that  
 CC are currently used.  
 XX SQ Sequence 49 AA;  
 OY Query Match 65.7%; Score 71; DB 21; Length 49;  
 Best Local Similarity 80.0%; Pred. No. 0.0099;  
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Db 22 ppgssghwsyglrpg 36  
 OY 3 PPSSEHWSYGLRPG 17  
 Db 22 ppgssghwsyglrpg 36  
 RESULT 13  
 Y58135  
 ID Y58135 standard; Protein; 49 AA.  
 XX AC Y58135;  
 XX DT 07-MAR-2000 (first entry)  
 DT

XX XX GnRH analogue multimer, containing four copies of the GnRH analogue.  
 XX DE Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;  
 KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;  
 KW non-androgenic; steroid; reduction; weight gain; muscle distribution;  
 KW fat distribution; male pattern; boar taint; flavour; impairment;  
 KW reliable; immunocastration; meat production.  
 XX OS Synthetic.  
 OS Mammalia.  
 XX MO9956771-A2.  
 XX PN 11-NOV-1999.  
 XX PD 05-MAY-1999; 99WO-CA00360.  
 XX PF 05-MAY-1999; 98US-0084217.  
 XX PR 05-MAY-1998; 98US-0084217.  
 XX PA (BIOS-) BIOSTAR INC.  
 XX PI Manns JG, Acres SD, Harland R;  
 XX PI WPI; 2000-062125/05.  
 XX DR N-PSDB; 246402.  
 XX DR  
 XX PT Production of uncastrated male food animals using vaccines -  
 PT  
 XX PS Example 1; Fig 2B; 87pp; English.  
 XX CC This sequence represents four copies of a gonadotropin  
 CC releasing hormone (GnRH) analogue, DNA encoding which was  
 CC used in the construction of a chimeric GnRH-leukotoxin (LKT)  
 CC fusion gene (Z46400). This fusion gene encodes a GnRH-LKT fusion  
 CC protein which may be used as a vaccine. The LKT portion of the protein  
 CC acts to enhance the immunogenicity of the GnRH portion. The invention  
 CC relates to a method of using two GnRH immunogen vaccines to produce  
 CC uncastrated male animals for meat production, one vaccination prior to  
 CC or during the fattening period to reduce circulating testosterone levels,  
 CC and the second vaccination about 2-8 weeks before slaughter to  
 CC substantially reduce androgenic and/or non-androgenic steroids. The  
 CC invention is used to produce food animals that exhibit the weight gain  
 CC and muscle/fat distribution of male animals without the problems  
 CC associated with male animals. Such problems include "boar taint", a  
 CC urine-like odour found in cooked meat of uncastrated pigs which is  
 CC caused by steroids stored in the tissues, and similar flavour  
 CC impairments in the meat of other intact male animals. The invention is  
 CC more reliable than prior art immunocastration techniques.  
 XX SQ Sequence 49 AA;  
 OY Query Match 65.7%; Score 71; DB 21; Length 49;  
 Best Local Similarity 80.0%; Pred. No. 0.0099;  
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Db 22 ppgssghwsyglrpg 36  
 OY 3 PPSSEHWSYGLRPG 17  
 Db 22 ppgssghwsyglrpg 36  
 RESULT 14  
 W03943  
 ID W03943 standard; Protein; 544 AA.  
 XX AC W03943;  
 XX DT 20-NOV-1996 (first entry)  
 XX DE LKT-GnRH protein fusion from pCB11.  
 XX DT Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;  
 KW

KW		fusion protein; immunogen; vaccine; fertility control;
KW		contraceptive; sterilisation.
OS		Chimeric Pasteurella haemolytica A1 strain B122;
OS		Chimeric synthetic.
XX		
FH	Key	Location/Qualifiers
FT	Domain	1..493
FT	/label= LKT	
FT	494..544	
FT	/label= GnRH_repeat_domain	
PN	M09624675-A1.	
PN		
PD	15-AUG-1996.	
PF	24-JAN-1996;	96WO-CA00049.
PR	10-FEB-1995;	95US-0387156.
PA	(UYSA-) UNIV SASKATCHEWAN.	
PI	Manns JG, Potter AA;	
DR	WPt: 1996-384447/38.	
N-P	SDB: T37177.	
PT	Gonadotropin-releasing hormone multimer fusion proteins - with leukotoxin polypeptide for increased immunogenicity, useful in antifertility vaccine prodn.	
PS	Claim 8; Fig 7A-7E; 87pp: English.	
CC	A chimeric protein (M03943) is composed of a fusion between a truncated leukotoxin (LKT-111) from Pasteurella haemolytica and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat sequence (see also M03944). It is the product of a chimeric gene (T37177) produced by deleting an approx. 1300 bp sequence from PCB113 (see also T37176) coding for amino acids 352-784 of LKT-352. Recombinant plasmid PCB111 (LKT 111:4 copy GnRH, ATCC 69748) was obtd. Escherichia coli transformants produced the chimeric protein, which is useful as a vaccine for fertility control, esp. immunological sterilisation of domestic or farm animals.	
SQ	Sequence 544 AA;	
Query Match	65.7%;	Score 71; DB 17; Length 544;
Best Local Similarity	80.0%;	Pred. No. 0.099;
Matches 12; Conservative	1;	Mismatches 2; Indels 0; Gaps 0;
OY	3 PPPSEHHWSYGRRP 17     :	
Db	515 pgssqmwsgylrpg 529	
RESULT 15		
W79570		
ID	W79570 standard; Protein: 544 AA.	
AC	W79570;	
DT	24-DEC-1998 (first entry)	
DE	LKT-GnRH chimeric protein.	
XX		
KW	Chimera; PCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;	
KW	Cytotoxic activity; antigen presentation; immune response; vaccine;	
XX	tumour.	
XS		
SS	Synthetic.	
IX		

Query Match	Best Local Similarity	Score 71:	DB 19:	Length 544:
Matches 12:	Conservative	1:	Mismatches	2:
Indels	0:	Gaps	0	

  

Sequence	544 AA:
MO9806848-A1.	
19-FEB-1998.	
08-AUG-1997;	97WO-CA00559.
09-AUG-1996;	96US-0694865.
(UYSA-) UNIV SASKATCHEWAN.	
Manns JG, Potter AA:	
WPI; 1998-159540/14.	
N-PSDB; V61532.	
Chimeric protein of leukotoxin and gonadotropin releasing hormone -	
useful for, e.g. preparation of vaccines for reduction of incidence	
of mammary tumours in mammals	
Disclosure: Figure 7.1-5; 118pp; English.	
The present sequence represents the LKT-GNRH chimeric protein from	
PCB11. This plasmid contains the LKT 111 polypeptide fused to	
four copies of the GNRH peptide. This chimera lacks cytotoxic activity	
which enables there to be an increase in antigen presentation and thus an	
optimal immune response. The removal of this region also enables the	
truncated LKT to be expressed at much higher levels and allows the amount	
of antigen administered to be reduced. This chimeric protein comprises a	
leukotoxin polypeptide, several trimers, and a GNRH sequence. The	
chimeric protein can be used as a vaccine to help reduce the incidence of	
mammary tumours in a mammalian individual.	

Search completed: March 2, 2001, 10:53:46  
Job time: 911 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:59:02 ; Search time 201.99 Seconds  
(without alignments)  
12.920 Million cell updates/sec

Title: US-09-306-689-3

Perfect score: 108

Sequence: 1 CPPPSEHWSYGLRPG 17

Scoring table: BLOSUM62

Searched: 957798 seqs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents-AA:\*  
1: /cgn2\_6/ptodata/2/paa/PCURS.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082.COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087.COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088.COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089.COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090.COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091.COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092.COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093.COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094.COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095.COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096.COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097.COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098.COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099.COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100.COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101.COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US102.COMB.pep:\*  
27: /cgn2\_6/ptodata/2/paa/US103.COMB.pep:\*  
28: /cgn2\_6/ptodata/2/paa/US104.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	17	US-09-305-924-3	Sequence 3, Appli
2	108	100.0	17	US-09-306-689-3	Sequence 3, Appli
3	102	94.4	17	PCT-US95-01225-5	Sequence 5, Appli
4	102	94.4	17	US-08-478-546B-5	Sequence 5, Appli
5	82.5	76.4	16	PCT-US95-01225-7	Sequence 7, Appli

6	82.5	76.4	16	8	US-08-478-546B-7	Sequence 7, Appli
7	81	75.0	34	1	PCT-US99-13960-56	Sequence 56, Appli
8	81	75.0	34	1	PCT-US99-13975B-122	Sequence 122, Appli
9	81	75.0	34	15	US-09-100-414-56	Sequence 56, Appli
10	81	75.0	34	15	US-09-100-414A-82	Sequence 82, Appli
11	81	75.0	34	17	US-09-303-323-82	Sequence 82, Appli
12	81	75.0	34	27	US-09-701-588-122	Sequence 122, Appli
13	81	75.0	35	1	PCT-US99-13960-55	Sequence 55, Appli
14	81	75.0	35	1	PCT-US99-13975B-120	Sequence 120, Appli
15	81	75.0	35	15	PCT-US99-13975B-121	Sequence 121, Appli
16	81	75.0	35	15	US-09-100-414-55	Sequence 55, Appli
17	81	75.0	35	15	US-09-100-414A-80	Sequence 80, Appli
18	81	75.0	35	15	US-09-100-414A-81	Sequence 81, Appli
19	81	75.0	35	17	US-09-303-323-80	Sequence 80, Appli
20	81	75.0	35	17	US-09-303-323-81	Sequence 81, Appli
21	81	75.0	35	27	US-09-701-588-120	Sequence 120, Appli
22	81	75.0	35	27	US-09-701-588-121	Sequence 121, Appli
23	72.5	67.1	397	19	US-09-506-078-31	Sequence 31, Appli
24	72.5	67.1	397	22	US-60-120-454-31	Sequence 31, Appli
25	72.5	67.1	398	22	US-60-120-454-25	Sequence 25, Appli
26	72.5	67.1	398	22	US-60-120-454-27	Sequence 27, Appli
27	72.5	67.1	442	19	US-09-506-078-27	Sequence 27, Appli
28	72.5	67.1	442	22	US-60-120-454-27	Sequence 27, Appli
29	71	65.7	49	14	US-09-019-010-4	Sequence 4, Appli
30	71	65.7	49	16	US-09-249-447A-7	Sequence 7, Appli
31	71	65.7	49	17	US-09-305-924-11	Sequence 11, Appli
32	71	65.7	49	17	US-09-306-689-11	Sequence 11, Appli
33	71	65.7	49	17	US-09-383-912-4	Sequence 4, Appli
34	71	65.7	544	17	US-09-383-912-10	Sequence 10, Appli
35	71	65.7	695	17	US-09-305-924-13	Sequence 13, Appli
36	71	65.7	695	17	US-09-306-689-13	Sequence 13, Appli
37	71	65.7	699	17	US-09-383-912-16	Sequence 16, Appli
38	71	65.7	977	17	US-09-383-912-8	Sequence 8, Appli
39	69	63.9	49	22	US-60-192-739-3463	Sequence 3463, Ap
40	68	63.0	14	5	US-08-160-882-24	Sequence 24, Appli
41	68	63.0	24	5	US-08-160-882-43	Sequence 43, Appli
42	68	63.0	44	5	US-08-160-882-45	Sequence 45, Appli
43	68	63.0	84	5	US-08-160-882-47	Sequence 47, Appli
44	68	63.0	90	13	US-08-912-314A-19	Sequence 19, Appli
45	66	61.1	17	5	US-08-160-882-16	Sequence 16, Appli

#### ALIGNMENTS

RESULT 1  
US-09-305-924-3  
; Sequence 3, Application US/09305924A  
; GENERAL INFORMATION:  
; APPLICANT: Jack G. Manns  
; APPLICANT: Stephen D. Acres  
; APPLICANT: Richard Hatland  
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION  
; FILE REFERENCE: 9001-0048  
; CURRENT APPLICATION NUMBER: US/09/305,924A  
; EARLIER FILING DATE: 1999-05-05  
; EARLIER APPLICATION NUMBER: US 60/084,217  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: GNRH  
US-09-305-924-3

Query Match 100.0%; Score 108; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.2e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CPPPSEHWSYGLRPG 17  
|||||

Db 1 CPPPSEHWSYGLRPG 17

RESULT 2

US-09-306-689-3

Sequence 3, Application US/09306689B

GENERAL INFORMATION:

APPLICANT: Robbins, Sarah C.

TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN

TITLE OF INVENTION: ANIMALS

FILE REFERENCE: 9001-0047

CURRENT APPLICATION NUMBER: US/09/306,689B

CURRENT FILING DATE: 1999-05-06

EARLIER APPLICATION NUMBER: US 60/088,024

EARLIER FILING DATE: 1998-06-04

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 17

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Sequence

US-09-306-689-3

Query Match 100.0%; Score 108; DB 17; Length 17;

Best Local Similarity 100.0%; Pred. No. 8,2e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 CPPPSEHWSYGLRPG 17

Db 1 CPPPSEHWSYGLRPG 17

RESULT 3

PCT-US95-01225-5

Sequence 5, Application PC/RUS9501225

GENERAL INFORMATION:

APPLICANT: Grimes, Stephen

APPLICANT: Scibienski, Robert

TITLE OF INVENTION: Immunogens Against Gonadotropin

TITLE OF INVENTION: Releasing Hormone

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dimitrios T. Drivas, Esq.

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/01225

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Drivas Esq., Dimitrios T.

REGISTRATION NUMBER: 32,218

REFERENCE/DOCKET NUMBER: 1102865-300

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-819-8286

TELEFAX: 212-354-8113

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: YES

FEATURE:

NAME/KEY: Region

LOCATION: 1..17

OTHER INFORMATION: /note= "spacer"

FEATURE:

NAME/KEY: Region

LOCATION: 8..17

OTHER INFORMATION: /note= "immunomimic"

FEATURE:

NAME/KEY: Modified-site

LOCATION: 17

OTHER INFORMATION: /label= glyNH2

OTHER INFORMATION: /note= "glycinamide"

PCT-US95-01225-5

Query Match 94.4%; Score 102; DB 1; Length 17;

Best Local Similarity 100.0%; Pred. No. 4,5e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 CPPPSEHWSYGLRP 16

Db 1 CPPPSEHWSYGLRP 16

RESULT 4

US-08-478-546B-5

Sequence 5, Application US/08478546B

GENERAL INFORMATION:

APPLICANT: Grimes, Stephen

APPLICANT: Scibienski, Robert

TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent

TITLE OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormo

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dimitrios T. Drivas, Esq.

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,546B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/188,223

FILING DATE: 27-JAN-1994

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Drivas Esq., Dimitrios T.

REGISTRATION NUMBER: 32,218

REFERENCE/DOCKET NUMBER: 1102865-300

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-819-8286

TELEFAX: 212-354-8113

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Region

LOCATION: 1..7

OTHER INFORMATION: /note= "spacer"



FEATURE:  
NAME/KEY: Region  
LOCATION: 8..17  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Modified site  
LOCATION: 17  
OTHER INFORMATION: /note= Xaa  
OTHER INFORMATION: /note= "amidated glycine"  
US-08-478-546B-5

Query Match 94.4%; Score 102; DB 8; Length 17;  
Best Local Similarity 100.0%; Pred. No. 4.5e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CPPPSSEHWSYGLRP 16  
DB 1 CPPPSSEHWSYGLRP 16

RESULT 5  
PCT-US95-01225-7  
Sequence 7, Application PC/TUS9501225  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01225  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..6  
OTHER INFORMATION: /note= "spacer"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 7..16  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 16  
OTHER INFORMATION: /label= GlyNH2  
OTHER INFORMATION: /note= "glycinamide"

PCT-US95-01225-7

Query Match 76.4%; Score 82.5; DB 1; Length 16;  
Best Local Similarity 87.5%; Pred. No. 0.0011;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 CPPPSSEHWSYGLRP 16  
DB 1 CPPPP-REHWSYGLRP 15

RESULT 6  
US-08-478-546B-7  
Sequence 7, Application US/08478546B  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent  
Tumors with Immunogens against Gonadotropin Releasing Hormo  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,546B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/188,223  
FILING DATE: 27-JAN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218  
REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..6  
OTHER INFORMATION: /note= "spacer"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 7..16  
OTHER INFORMATION: /note= "immunomimic"  
FEATURE:  
NAME/KEY: Modified site  
LOCATION: 16  
OTHER INFORMATION: /note= Xaa  
OTHER INFORMATION: /note= "amidated glycine"  
US-08-478-546B-7

Query Match 76.4%; Score 82.5; DB 8; Length 16;  
Best Local Similarity 87.5%; Pred. No. 0.0011;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 CPPPSEHWSYGLRP 16  
| | | | | | | | | |  
Db 1 CPPP-REHWSYGLRP 15

RESULT 7  
PCT-US99-13960-56  
; Sequence 56, Application PC/TUS9913960  
; GENERAL INFORMATION:  
; APPLICANT: UNITED BIOMEDICAL INC.  
; TITLE OF INVENTION: ARTIFICIAL T HELPER CELL EPITOPES  
; TITLE OF INVENTION: AS IMMUNE STIMULATORS FOR SYNTHETIC PEPTIDE IMMUNOGENS INCLUDED  
; TITLE OF INVENTION: IMMUNOGENIC LHRH PEPTIDES  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Flinnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC Windows  
; SOFTWARE: Word 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US99/13960  
; FILING DATE: 21-JUNE-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/100,414  
; FILING DATE: 20-JUNE-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4157PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US99-13960-56

Query Match 75.0%; Score 81; DB 1; Length 34;  
Best Local Similarity 81.2%; Pred. No. 0.0033;  
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PPPPSEHWSYGLRP 17  
| | | | | | | | | |  
Db 19 PPXPPEHWSYGLRP 34

RESULT 8  
PCT-US99-13975B-122  
; Sequence 122, Application PC/TUS9913975B  
; GENERAL INFORMATION:  
; APPLICANT: UNITED BIOMEDICAL INC., ET AL.  
; TITLE OF INVENTION: ARTIFICIAL T HELPER CELL  
; TITLE OF INVENTION: EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC  
; TITLE OF INVENTION: PEPTIDE IMMUNOGENS  
; NUMBER OF SEQUENCES: 151  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Flinnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA

ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC Windows  
; SOFTWARE: Word 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US99/13975B  
; FILING DATE: 21-JUNE-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/100,412  
; FILING DATE: 20-JUNE-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4158PC1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 122:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US99-13975B-122

Query Match 75.0%; Score 81; DB 1; Length 34;  
Best Local Similarity 81.2%; Pred. No. 0.0033;  
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PPPPSEHWSYGLRP 17  
| | | | | | | | | |  
Db 19 PPXPPEHWSYGLRP 34

RESULT 9  
US-09-100-414-56  
; Sequence 56, Application US/09100414  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang YI  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; TITLE OF INVENTION: IMMUNOGENS  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Flinnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC Windows  
; SOFTWARE: Word 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/100,414  
; FILING DATE: 20-JUNE-1998  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4157  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34 amino acids  
; TYPE: amino acid

ADDRESSEE: Morgan & Finegan, L.L.P.

REFERENCE/DOCKET NUMBER: 1151-158US1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 122:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 122:  
US-09-701-588-122

Query Match 75.0%; Score 81; DB 27; Length 34;  
Best Local Similarity 81.2%; Pred. No. 0.0033;  
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PPPSSEHWSYGLRPG 17  
||| |||||  
Db 19 PPXPPEHWSYGLRPG 34

RESULT 13  
PCT-US99-13960-55  
Sequence 55, Application PC/TUS9913960  
GENERAL INFORMATION:  
APPLICANT: UNITED BIOMEDICAL INC.  
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL EPITOPES  
TITLE OF INVENTION: AS IMMUNE STIMULATORS FOR SYNTHETIC PEPTIDE IMMUNOGENS INCLUDED  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US99/13960  
FILING DATE: 21-JUNE-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria C.H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 4  
OTHER INFORMATION: /note= "Ser or  
OTHER INFORMATION: Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 7  
OTHER INFORMATION: /note= "Lys or  
OTHER INFORMATION: Arg"  
FEATURE:

NAME/KEY: Modified-site  
LOCATION: 8  
OTHER INFORMATION: /note= "Gly or  
OTHER INFORMATION: Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 12  
OTHER INFORMATION: /note= "His or  
OTHER INFORMATION: Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 13  
OTHER INFORMATION: /note= "Lys or  
OTHER INFORMATION: Arg"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 16  
OTHER INFORMATION: /note= "Gly or  
OTHER INFORMATION: Thr"  
PCT-US99-13960-55

Query Match 75.0%; Score 81; DB 1; Length 35;  
Best Local Similarity 81.2%; Pred. No. 0.0034;  
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PPPSSEHWSYGLRPG 17  
||| |||||  
Db 20 PPXPPEHWSYGLRPG 35

RESULT 14  
PCT-US99-13975B-120  
Sequence 120, Application PC/TUS9913975B  
GENERAL INFORMATION:  
APPLICANT: UNITED BIOMEDICAL INC., ET AL.  
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL  
TITLE OF INVENTION: EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC  
NUMBER OF SEQUENCES: 151  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US99/13975B  
FILING DATE: 21-JUNE-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,412  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4158PC1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 120:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US99-13975B-120

Query Match 75.0%; Score 81; DB 1; Length 35;  
 Best Local Similarity 81.2%; Pred. NO. 0.0034;  
 Matches 13; Conservative 0; Mismatches 3; Indels 0;

OY 2 PPPSEHWSYGLRPG 17  
 ||| |||||  
 Db 20 PPXPPEHWSYGLRPG 35

## RESULT 15

PCT-US99-13975B-121  
 ; Sequence 121, Application PC/TUS9913975B  
 ; GENERAL INFORMATION:  
 ; APPLICANT: UNITED BIOMEDICAL INC., ET AL.  
 ; TITLE OF INVENTION: ARTIFICIAL T HELPER CELL  
 ; TITLE OF INVENTION: EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC  
 ; TITLE OF INVENTION: PEPTIDE IMMUNOGENS  
 ; NUMBER OF SEQUENCES: 151  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morgan & Finnegan, L.L.P.  
 ; STREET: 345 Park Avenue  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10154-0054  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC Windows  
 ; SOFTWARE: Word 97  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US99/13975B  
 ; FILING DATE: 21-JUNE-1999  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/100,412  
 ; FILING DATE: 20-JUNE-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Maria H. Lin  
 ; REGISTRATION NUMBER: 29,323  
 ; REFERENCE/DOCKET NUMBER: 1151-4158P01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-758-4800  
 ; TELEFAX: 212-751-6849  
 ; INFORMATION FOR SEQ ID NO: 121:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 35 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; PCT-US99-13975B-121

Query Match 75.0%; Score 81; DB 1; Length 35;  
 Best Local Similarity 81.2%; Pred. NO. 0.0034;  
 Matches 13; Conservative 0; Mismatches 3; Indels 0;

OY 2 PPPSEHWSYGLRPG 17  
 ||| |||||  
 Db 20 PPXPPEHWSYGLRPG 35

Search completed: March 2, 2001, 10:59:03  
 Job time: 373 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:46 ; Search time 66.16 Seconds  
(without alignments)  
8.786 Million cell updates/sec

Title: US-09-306-689-4

Perfect score: 104

Sequence: 1 XHMSYGLRPSSPPPC 17

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

268485

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_36:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	99.0	17	16 R78283	GnRH immunomimic a
2	103	99.0	17	21 Y58140	Gonadotropin relea
3	83.5	80.3	16	16 R78285	GnRH immunomimic a
4	83.5	80.3	16	21 Y58141	Gonadotropin relea
5	59.6	59.6	14	15 R44953	E. coli fimbriae H
6	59.6	59.6	49	17 W03944	GnRH 4-repeat sequ
7	59.6	59.6	49	19 W79567	GnRH-2. Synthetic
8	59.6	59.6	49	19 W61542	Peptide hormone Gn
9	59.6	59.6	49	21 Y58363	Four-copy gonadotr
10	59.6	59.6	49	21 Y58135	GnRH analogue mult
11	59.6	59.6	257	12 R11178	Plasmid pBTA735-en
12	59.6	59.6	257	12 R11183	Plasmid pBTA740-en

13	62	59.6	544	17 W03943	LKT-GnRH protein f
14	62	59.6	544	19 W79570	LKT-GnRH chimeric
15	62	59.6	695	19 W79573	LKT-GnRH chimeric
16	62	59.6	695	21 Y58361	Leukotoxin/gonadot
17	62	59.6	695	21 Y58133	Gonadotropin relea
18	62	59.6	977	17 W03942	LKT-GnRH protein f
19	62	59.6	977	17 W79569	LKT-GnRH chimeric
20	59	56.7	12	15 R44954	E. coli fimbriae H
21	59	56.7	18	21 Y89763	Core polypeptide f
22	59	56.7	253	12 R11180	Plasmid pBTA734-en
23	59	56.7	253	12 R11181	Plasmid pBTA733-en
24	59	56.7	253	12 R11182	Plasmid pBTA731-en
25	59	56.7	263	12 R11185	Plasmid pBTA70-en
26	59	56.7	283	12 R11186	Plasmid pBTA862-en
27	59	56.7	323	12 R11187	Plasmid pBTA859-en
28	58	55.8	9	20 W94891	LHRH peptide fragm
29	58	55.8	10	2 P10097	Sequence of lutein
30	58	55.8	10	2 P10411	Luteinising Hormon
31	58	55.8	10	2 P10416	Luteinising Hormon
32	58	55.8	10	6 P50222	Gonadotropin rele
33	58	55.8	10	7 P60127	Gonadotropin rele
34	58	55.8	10	7 P61403	Gonadotropin relea
35	58	55.8	10	7 P60576	Novel decapeptide
36	58	55.8	10	8 P70922	Luteinising hormon
37	58	55.8	10	10 P90630	Sequence of lutein
38	58	55.8	10	12 R15713	Peptide #1 with ho
39	58	55.8	10	13 R26819	LH releasing hormo
40	58	55.8	10	15 R62689	LHRH hapten for at
41	58	55.8	10	16 R91197	LHRH peptide. Syn
42	58	55.8	10	16 R86845	Gonadotropin relea
43	58	55.8	10	16 R75152	Gonadotropin relea
44	58	55.8	10	17 W65201	Luteinising hormon
45	58	55.8	10	17 Y55061	Luteinising hormon

#### ALIGNMENTS

RESULT 1	
ID R78283	standard; peptide: 17 AA.
XX	
AC R78283;	
XX	
DT 13-MAR-1996	(first entry)
XX	
DE GnRH immunomimic and spacer (GnRH(1-10)-Ser10).	
XX	
KW Immunomimetic; gonadotropin releasing hormone; GnRH; LHRH; vaccine;	
KW luteinising hormone releasing hormone; spacer; immunomimic; uterine;	
KW diphtheria toxin; DT; gynaecological; endometriosis; uterine fibroids;	
KW gonadal steroid hormone associated dependent disease; gonadotropin;	
KW immunological contraception; mammal; breast; cancer; prostate;	
KW benign prostatic hypertrophy.	
XX	
OS Chimeric - Homo sapiens.	
OS	
XX	
OS Chimeric - Synthetic.	
XX	
FH Key	Location/Qualifiers
FT Modified-site	1
FT	/label= OTHER
FT	/note= "pyroglutamic acid"
FT	1..10
FT	/note= "GnRH"
FT	11..17
FT	/note= "spacer"
XX	
PN W09520600-A1.	
XX	
PD 03-AUG-1995.	
XX	
PF 26-JAN-1995;	95WO-US01225.
XX	





XX Grimes S, Scibienski R;  
 XX WPI; 1995-275410/36.  
 DR  
 XX  
 XX New anti-gonadotropin releasing hormone immunogenic composition(s)  
 PT - used for treating gonadotropin and gonadal steroid hormone  
 PT dependent disease(s) and providing contraception.  
 PS Claim 1, Page 29; 39pp; English.  
 XX  
 XX The sequences given in R78282-85 represent immunomimetics to  
 CC gonadotropin releasing hormone (GnRH). These peptides comprise the  
 CC wild type GnRH sequence and a spacer attached to either the N- or C-  
 CC terminal. The spacer molecules serve as a link through which the  
 CC immunomimic is attached to an immunological carrier such as  
 CC diphtheria toxoid (DT) and also affects the immune response generated  
 CC by the vaccinated mammal against the immunomimic. Compositions  
 CC comprising these peptides may be used for treating a mammal for  
 CC gonadotropin and gonadal steroid hormone associated dependent  
 CC disease or for providing immunological contraception in mammals.  
 CC They can also be used for treating breast cancer, uterine and other  
 CC gynaecological cancers, endometriosis, uterine fibroids, prostate  
 CC cancer, or benign prostatic hypertrophy.  
 CC  
 SQ Sequence 16 AA;

Query Match 80.3%; Score 83.5; DB 16; Length 16;  
 Best Local Similarity 87.5%; Pred. No. 0.00017;  
 Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 HWSYGLRPGSSPPPC 17  
 |||||  
 Db 2 hwsyglrpg-rppppc 16

RESULT 4  
 Y58141  
 ID Y58141 standard; peptide; 16 AA.  
 XX  
 AC Y58141;

DT 07-MAR-2000 (first entry)

DE Gonadotropin releasing hormone (GnRH) peptide analogue 5.

XX Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;  
 KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;  
 KW non-androgenic; steroid; reduction; weight gain; muscle distribution;  
 KW fat distribution; male pattern; boar taint; flavour; impairment;  
 KW reliable; immunocastration; meat production.  
 XX

OS Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "Pyroglutamic acid"

W09356771-A2.

PD 11-NOV-1999.

PE 05-MAY-1999; 99WO-CA00360.

PR 05-MAY-1998; 98US-0084217.

PA (BIOS-) BIOSSTAR INC.

PI Manns JG, Acres SD, Harland R;

DR WPI; 2000-062125/05.

XX

PT Production of uncastrated male food animals using vaccines -  
 XX  
 XX Disclosure; Page 11; 87pp; English.  
 PS  
 XX  
 XX Sequences Y58136-Y58141 represent gonadotropin releasing hormone  
 CC (GnRH) analogues which may be used as an alternative to sequence  
 CC Y58135 in embodiments of the present invention. The invention  
 CC relates to a method of using two GnRH immunogen vaccines to produce  
 CC uncastrated male animals for meat production, one vaccination prior to  
 CC or during the fattening period to reduce circulating testosterone levels,  
 CC and the second vaccination about 2-8 weeks before slaughter to  
 CC substantially reduce androgenic and/or non-androgenic steroids. The  
 CC invention is used to produce food animals that exhibit the weight gain  
 CC and muscle/fat distribution of male animals without the problems  
 CC associated with male animals. Such problems include "boar taint", a  
 CC urine-like odour found in cooked meat of uncastrated pigs which is  
 CC caused by steroids stored in the tissues, and similar flavour  
 CC impairments in the meat of other intact male animals. The invention is  
 CC more reliable than prior art immunocastration techniques.  
 CC  
 SQ Sequence 16 AA;

Query Match 80.3%; Score 83.5; DB 21; Length 16;  
 Best Local Similarity 87.5%; Pred. No. 0.00017;  
 Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 HWSYGLRPGSSPPPC 17  
 |||||  
 Db 2 hwsyglrpg-rppppc 16

RESULT 5  
 R44953  
 ID R44953 standard; Protein; 14 AA.  
 XX  
 AC R44953;

DT 07-JUL-1994 (first entry)

DE E. coli fimbriae HR4 GnRH insert linker.

XX Gonadotropin releasing hormone; immunogenic carrier system;  
 KW immune response; immunisation; vaccine; contraceptive agent;  
 KW treatment; sexual hyperactivity; cancer.  
 XX

OS Synthetic.

XX Key Location/Qualifiers  
 FH Peptide 2..11  
 FT /note= "decapeptide GnRH"

EP578293-A.

PD 12-JAN-1994.

PE 15-JUN-1993; 93EP-0201712.

PR 18-JUN-1992; 92EP-0201775.

PA (ALKU) AKZO NV.

PI Gielen JT, Hoekstra WPM, Van Der Zee A, Van Die IM;

DR WPI; 1994-010165/02.

N-PsDB; Q54392.

PT Immunogenic carrier system comprising Escherichia coli P-fimbriae  
 PT - used to elicit immune response against gonadotropin releasing  
 hormone  
 XX

Example 1; Fig 2; 34pp; English.

XX

CC The sequence is that of an insert contg. the decapeptide gonadotropin  
CC releasing hormone (GnRH) located in the major subunit of P-fimbrinae  
CC with serotype III at a position corresponding to a position in  
CC hypervariable region 4 of the wild type major subunit. This forms  
CC part of an immunogenic carrier system capable of eliciting an immune  
CC response against GnRH.

XX Sequence 14 AA;

Query Match 59.6%; Score 62; DB 15; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.066; Mismatches 0; Gaps 0;

QY 2 HWSYGLRPGS 11  
| | | | | | | | | |  
DB 3 hwsyglrpgs 12

RESULT 6

W03944 W03944 standard; Protein; 49 AA.

XX AC W03944;

XX DT 20-NOV-1996 (first entry)

XX DE GnRH 4-repeat sequence.

XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;  
XX fusion protein; immunogen; vaccine; fertility control;  
XX KW contraceptive; sterilisation; PCB113; PCB11.  
XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Peptide 1..10

FT Peptide /label= GnRH

FT Peptide 11..13

FT Peptide /label= Spacer

FT Peptide 14..23

FT Peptide /label= GnRH

FT Peptide 24..26

FT Peptide /label= Spacer

FT Peptide 27..36

FT Peptide /label= GnRH

FT Peptide 37..39

FT Peptide /label= Spacer

FT Peptide 40..49

FT Peptide /label= GnRH

XX W09624675-A1.

XX PN 15-AUG-1996.

XX PD 24-JAN-1996; 96WO-CA00049.

XX PR 10-FEB-1995; 95US-0387156.

XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Manns JG, Potter AA;

XX DR WPI: 1996-384447/38.

XX N-PSDB: T37178.

XX Gonadotropin-releasing hormone multimer fusion proteins - with  
XX P leucoxin polypeptide for increased immunogenicity, useful in  
XX PT antifertility vaccine prodn.  
XX PS Example 2; Fig 1B; 87pp; English.

CC A synthetic DNA sequence (T37178) codes for a gonadotropin

CC releasing hormone (GnRH) tetramer (W03944), in which the 4 GnRH  
CC repeat units are separated by spacers designed to increase  
CC immunogenicity. The DNA sequence was incorporated into vector  
CC pAA352 (ATCC 68283), which contains a truncated leukotoxin  
CC gene (LKT 352) derived from Pasteurella haemolytica, to  
CC give plasmid PCB113 (T37176). Escherichia coli transformants  
CC produce an LKT-GnRH fusion protein (see also W03942 and W03943)  
CC useful as a vaccine for fertility control of domestic or farm animals.

XX Sequence 49 AA;

Query Match 59.6%; Score 62; DB 17; Length 49;

Best Local Similarity 100.0%; Pred. No. 0.21; Mismatches 0; Gaps 0;

QY 2 HWSYGLRPGS 11  
| | | | | | | | | |  
DB 2 hwsyglrpgs 11

RESULT 7

W79567 W79567 standard; Protein; 49 AA.

XX AC W79567;

XX DT 24-DEC-1998 (first entry)

XX DE GnRH-2.

XX Gonadotropin releasing hormone; GnRH; decapeptide; hypothalamus; spacer;  
XX LH; luteinising hormone; FSH; follicle stimulating hormone; vertebrate;  
XX KW pyrogliu; chimera; leukotoxin polypeptide; multimer; vaccine; tumour;  
XX immunogenic.

XX OS Synthetic.

XX PN W09806848-A1.

XX PD 19-FEB-1998.

XX PF 08-AUG-1997; 97WO-CA00559.

XX PR 09-AUG-1996; 96US-0694865.

XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Manns JG, Potter AA;

XX DR WPI: 1998-159540/14.

XX N-PSDB: V61529.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -  
XX PT useful for, e.g. preparation of vaccines for reduction of incidence  
XX of mammary tumours in mammals  
XX PS Disclosure; Figure 1B; 118pp; English.

XX The present sequence represents a recombinantly produced or chemically  
XX synthesised gonadotropin releasing hormone-2 (GnRH-2) polypeptide, which  
XX contains four copies of the GnRH decapeptide (V61528) and triplet amino  
XX acid spacers between each of these sequences. This decapeptide is  
XX secreted naturally by the hypothalamus which controls release of both the  
XX luteinising hormone (LH) and the follicle stimulating hormone (FSH) in  
XX vertebrates. This sequence, as compared to the native peptide, has been  
XX found to have an N-terminal Gln rather than a pyroglu residue, and also  
XX contains substitutions at amino acid residues 15 and 41, whereby His is  
XX replaced by Asp. This produces an alternating multimeric GnRH sequence  
XX which is highly immunogenic that can be used in the construction of a  
XX chimeric protein that comprises a leukotoxin polypeptide, several  
XX multimers, and the GnRH sequence. The chimeric protein can be used as a  
XX vaccine to help reduce the incidence of mammary tumours in a mammalian

CC Individual.  
 XX  
 SO Sequence 49 AA;

Query Match 59.6%; Score 62; DB 19; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HNSYGLRPGS 11  
 |||||  
 Db 2 hwsyglrpgs 11

RESULT 8  
 W61542  
 ID W61542 standard; Protein; 49 AA.

AC W61542;

DT 27-OCT-1998 (first entry)

DE Peptide hormone GnRH-2 decapeptide (4 copies) fragment.

XX GnRH; gonadotropin releasing hormone peptide hormone; leukotoxin;  
 XX immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;  
 KM immune response; hormone receptor; cancerous cell; domestic animal;  
 KM porcine; bovine; luteinizing hormone; follicle stimulating hormone;  
 KM immunocastrate.

OS Synthetic.

PN WO9834639-A1.

PD 13-AUG-1998.

PF 04-FEB-1998; 98WO-CA00059.

PR 05-FEB-1997; 97US-0036883.

PA (BIOS-) BIOSTAR INC.

PI Acres SD, Harland R, Manns JG;

DR WPI: 1998-446952/38.

DR N-PSDB: V45190.

PT Immunisation against endogenous molecules by administering vaccine  
 PT to ear - useful to elicit efficient and uniform immune response  
 PT against e.g. gonadotropin releasing hormone to. Immunocastrate pigs  
 PT and cattle

PS Example 1; Fig 1B; 61pp; English.

CC This represents the amino acid sequence of the gonadotropin releasing  
 CC hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-  
 CC GnRH polypeptide gene fusions. This is used to exemplify the method of  
 CC invention of immunisation against endogenous molecules by administering  
 CC a vaccine which comprises an immunogen and a carrier to the ear of the  
 CC mammal. The method is useful for eliciting an efficient and uniform  
 CC immune response to block or suppress the activity of an endogenous  
 CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,  
 CC or to elicit an immune response against a targeted endogenous cell type  
 CC (e.g. a cancerous or otherwise diseased cell). It is especially useful  
 CC to reduce the levels of GnRH in domestic animals, especially in porcine  
 CC or bovine species. The use of GnRH immunogens in the vaccine reduces the  
 CC levels of luteinizing hormone and follicle stimulating hormone and helps  
 CC in immunocastrating the animal. Administration of vaccine compositions to  
 CC the ear instead of intramuscular administration into the neck increases  
 CC the efficiency of vaccination of mammals against endogenous immunogens,  
 CC and may increase uniformity of vaccine presentation since the ear is  
 CC relatively uniform from animal to animal.

SO Sequence 49 AA;

Query Match 59.6%; Score 62; DB 19; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HNSYGLRPGS 11  
 |||||  
 Db 2 hwsyglrpgs 11

RESULT 9  
 Y58363  
 ID Y58363 standard; Protein; 49 AA.

AC Y58363;

DT 27-MAR-2000 (first entry)

DE Four-copy gonadotropin-releasing hormone (GnRH) multimer.

XX GnRH multimer; gonadotropin-releasing hormone; immunosterilisation;  
 XX immunocastration; vaccine; feline; canine; equine; cervine; ds  
 OS Mammalia.  
 OS Synthetic.

PN WO962545-A2.

PD 09-DEC-1999.

PF 28-MAY-1999; 99WO-CA00493.

PR 04-JUN-1998; 98US-0088024.

PR 06-MAY-1999; 99US-0306689.

PA (BIOS-) BIOSTAR INC.

PI Robbins SC;

DR WPI: 2000-086857/07.

DR N-PSDB: Z55702.

PT Hormone immunogens, analogues or antibodies used to manufacture  
 PT vaccines for suppression of reproductive behavior and fertility in  
 PT vertebrates -  
 PS Claim 5; Fig 5B; 88pp; English.

CC This sequence represents a four-copy gonadotropin-releasing  
 CC hormone (GnRH) multimer, where the second and fourth GnRH sequence  
 CC have a His to Asp substitution at position 2 of the GnRH sequence.  
 CC The invention relates to GnRH immunogens, analogues or antibodies  
 CC that cross-react with endogenous GnRH of a vertebrate. A specifically  
 CC claimed immunogenic fusion protein (Y58361) comprises, in the N to  
 CC C-terminal direction, a synthetic peptide sequence (Y58364), an eight  
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer  
 CC sequence of Y58363), the LRT protein (which functions as a carrier  
 CC protein), and a second eight copy GnRH multimer. The fusion protein may  
 CC be used in a vaccine composition for prepubertal administration to a  
 CC vertebrate subject to result in prolonged suppression of reproductive  
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are  
 CC used to manufacture a composition or vaccine for immunosterilisation or  
 CC immunocastration of feline, canine, equine or cervine subjects.  
 CC The vaccines are used to suppress reproductive behaviour and/or  
 CC fertility for at least 10 months. The prepubertal administration  
 CC results in a prolonged, long-term suppression of testicular development  
 CC and/or function in males, or a prolonged, long-term suppression of  
 CC ovarian development and/or function in females. The methods provide a  
 CC viable and desirable alternative to surgical forms of sterilisation that  
 CC are currently used.

SQ Sequence 49 AA:

Query Match  
Best Local Similarity 100.0%; Score 62; DB 21; Length 49;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11  
|||  
Db 2 hwsyglrpgs 11

RESULT 10  
Y58135 Y58135 standard; Protein; 49 AA.

XX AC Y58135;

XX DT 07-MAR-2000 (first entry)

XX DE GnRH analogue multimer, containing four copies of the GnRH analogue.

XX KW Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;  
XX KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;  
XX KW non-androgenic; steroid; reduction; weight gain; muscle distribution;  
XX KW fat distribution; male pattern; boar taint; flavour; impairment;  
XX KW reliable; immunocastration; meat production.

XX OS Synthetic.  
XX OS Mammalia.

XX PN WO9956771-A2.

XX PD 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00360.

XX PR 05-MAY-1998; 98US-0084217.

XX PA (BIOS-) BIOSTAR INC.

XX PI Manns JG, Acres SD, Harland R;

XX DR WPI: 2000-062125/05.

XX DR N-PSDB; 246402.

XX PT Production of uncastrated male food animals using vaccines -

XX PS Example 1; Fig 2B; 87pp; English.

XX CC This sequence represents four copies of a gonadotropin  
XX CC releasing hormone (GnRH) analogue, DNA encoding which was  
XX CC used in the construction of a chimeric GnRH-leukotoxin (LKT)  
XX CC fusion gene (Z46400). This fusion gene encodes a GnRH-LKT fusion  
XX CC protein which may be used as a vaccine. The LKT portion of the protein  
XX CC acts to enhance the immunogenicity of the GnRH portion. The invention  
XX CC relates to a method of using two GnRH immunogen vaccines to produce  
XX CC uncastrated male animals for meat production, one vaccination prior to  
XX CC or during the fattening period to reduce circulating testosterone levels,  
XX CC and the second vaccination about 2-8 weeks before slaughter to  
XX CC substantially reduce androgenic and/or non-androgenic steroids. The  
XX CC invention is used to produce food animals that exhibit the weight gain  
XX CC and muscle/fat distribution of male animals without the problems  
XX CC associated with male animals. Such problems include "boar taint", a  
XX CC urine-like odour found in cooked meat of uncastrated pigs which is  
XX CC caused by steroids stored in the tissues, and similar flavour  
XX CC impairments in the meat of other intact male animals. The invention is  
XX CC more reliable than prior art immunocastration techniques.

SQ Sequence 49 AA:

Query Match 59.6%; Score 62; DB 21; Length 49;

Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11  
|||  
Db 2 hwsyglrpgs 11

RESULT 11  
R11178 R11178 standard; Protein; 257 AA.

XX AC R11178;

XX DT 22-MAY-1991 (first entry)

XX DE Plasmid pBTAT735-encoded TrATp-LHRH fusion protein.

XX KW TrATp protein; Leutinizing hormone releasing hormone; fusion protein;  
XX KW immunological castration.

XX FH Key location/Qualifiers

XX FT Peptide 1..20 /label- TrATp signal

XX FT Peptide 103..112 /label- LHRH analogue

XX FT Protein 21..257 /label- TrATp-LHRH fusion

XX FN WO9102799-A.

XX PD 07-MAR-1991.

XX PF 24-AUG-1990; 90WO-AU00373.

XX PR 25-AUG-1989; 89AU-0005979.

XX PA (BIOT-) BIOTECHN AUST PTY L.

XX PI Russell-Jones GJ, Stewart AG, Tsonis CG;

XX DR WPI: 1991-087282/12.

XX DR N-PSDB; Q10999.

XX PT Fusion proteins comprising LHRH analogue and TrATp (analogue) -  
XX PT useful in vaccine for inhibition or control of reproduction in  
XX PT vertebrates, esp. domestic animals

XX PS Example 1; Fig 2 and 3; 53pp; English.

XX CC Plasmid pBTAT735 is a TrATp-LHRH analogue fusion in which the LHRH  
XX CC analogue has been inserted between amino acids 101 and 102 of TrATp  
XX CC (Ogata R.T. et al., (1982) J. Bacteriol. 151:819-827). The synthetic  
XX CC LHRH analogue was inserted via a linker which provides a unique new  
XX CC SmaI site located between codons such that the LHRH is inserted in  
XX CC frame. The fusion protein can be used in vaccines for the  
XX CC inhibition or control of reproduction in vertebrates, where the  
XX CC TrATp acts as an adjuvant.  
XX CC See also Q10995, Q10997-8, Q11000, Q11014-Q11021.

SQ Sequence 257 AA:

Query Match 59.6%; Score 62; DB 12; Length 257;  
Best Local Similarity 90.9%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSS 12  
|||  
Db 104 hwsyglrpgss 114

RESULT 12

R1183  
ID R1183 standard; Protein: 257 AA.  
XX  
AC R1183;  
XX  
DT 22-MAY-1991 (first entry)  
XX  
DE Plasmid pBTA740-encoded TratrP-LHRH fusion protein.  
XX  
KW Tratr protein; leutinizing hormone releasing hormone; fusion protein;  
KW Immunological castration.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..20  
FT /label= TratrP signal  
FT Peptide 238..247  
FT /label= LHRH analogue  
FT Protein 21..257  
FT /label= TratrP-LHRH fusion  
XX  
PN W09102799-A.  
XX  
PD 07-MAR-1991.  
XX  
PF 24-AUG-1990; 90MO-AU00373.  
XX  
PR 25-AUG-1989; 89AU-0005979.  
XX  
PA (BIOT-) BIOTECHN AUST PTY L.  
XX  
PI Russell-Jones GJ, Stewart AG, Tsouis CG;  
XX  
DR WPI; 1991-087282/12.  
DR N-PSDB; Q11017.  
XX  
XX Fusion proteins comprising LHRH analogue and TratrP (analogue) -  
PT useful in vaccine for inhibition or control of reproduction in  
PT vertebrates, esp. domestic animals  
XX  
PS Example 1; Fig 2 and 3; 53pp; English.  
XX  
XX Plasmid pBTA740 is a TratrP-LHRH analogue fusion in which the LHRH  
CC analogue has been inserted between amino acids 236 and 237 of TratrP  
CC (Ogata R.T. et al., (1982) J Bacteriol. 151:819-827). The synthetic  
CC LHRH analogue was inserted via a linker which provides a unique new  
CC SmaI site located between codons such that the LHRH is inserted in  
CC frame. The fusion protein can be used in vaccines  
CC for the inhibition or control of reproduction in vertebrates, where  
CC the TratrP acts as an adjuvant.  
XX  
CC See also Q10995, Q10997-Q11000, Q11014-6, Q11018-Q11021.  
XX  
SQ Sequence 257 AA;

Query Match 59.6%; Score 62; DB 12; Length 257;  
Best Local Similarity 90.9%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSS 12  
Db 239 hwsyglrpgss 249

RESULT 13  
ID W03943 standard; Protein: 544 AA.  
XX  
AC W03943;  
XX  
DT 20-NOV-1996 (first entry)  
XX  
DE LKT-GnRH protein fusion from PCB111.  
XX

KW Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;  
KW fusion protein; immunogen; vaccine; fertility control;  
KW contraceptive; sterilisation.  
XX  
XX OS Chimeric Pasteurella haemolytica A1 strain B122;  
XX Chimeric synthetic.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..493  
FT /label= LKT  
FT Domain 494..544  
FT /label= GnRH\_repeat\_domain  
XX  
PN W09624675-A1.  
XX  
PD 15-AUG-1996.  
XX  
PF 24-JAN-1996; 96MO-CA00049.  
XX  
PR 10-FEB-1995; 95US-0387156.  
XX  
PA (UYSA-) UNIV SASKATCHEWAN.  
XX  
PI Manns JG, Potter AA;  
XX  
DR WPI; 1996-384447/38.  
DR N-PSDB; T37177.  
XX  
XX Gonadotropin-releasing hormone multimer fusion proteins - with  
PT leukotoxin polypeptide for increased immunogenicity, useful in  
PT antifertility vaccine prodn.  
XX  
PS Claim 8; Fig 7A-7E; 87pp; English.  
XX  
XX A chimeric protein (W03943) is composed of a fusion between  
CC a truncated leukotoxin (LKT-111) from Pasteurella haemolytica  
CC and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat  
CC sequence (see also W03944). It is the product of a chimeric  
CC gene (T37177) produced by deleting an approx. 1300 bp sequence  
CC from PCB113 (see also T37176) coding for amino acids 352-784  
CC of LKT-352. Recombinant plasmid PCB111 (LKT 111:4 copy GnRH,  
CC ATCC 69748) was obtd. Escherichia coli transformants produced  
CC the chimeric protein, which is useful as a vaccine for fertility  
CC control, esp. immunological sterilisation of domestic or farm  
CC animals.  
XX  
SQ Sequence 544 AA;

Query Match 59.6%; Score 62; DB 17; Length 544;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11  
Db 495 hwsyglrpgs 504

RESULT 14  
ID W79570 standard; Protein: 544 AA.  
XX  
AC W79570;  
XX  
DT 24-DEC-1998 (first entry)  
XX  
DE LKT-GnRH chimeric protein.  
XX  
XX Chimera; PCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;  
KW cytotoxic activity; antigen presentation; immune response; vaccine;  
KW tumour.  
XX  
OS Synthetic.

XX MO9806848-A1.  
 PN 19-FEB-1998.  
 PD 08-AUG-1997; 97WO-CA00559.  
 PE 09-AUG-1996; 96US-0694865.  
 PR (UYSA-) UNIV SASKATCHEWAN.  
 PA Manns JG, Potter AA;  
 PI WPI; 1998-159540/14.  
 DR N-PSDB; V61532.  
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -  
 PT useful for, e.g. preparation of vaccines for reduction of incidence  
 of mammary tumours in mammals  
 PS Disclosure; Figure 7.1-5; 118pp; English.  
 XX The present sequence represents the LKT-GnRH chimeric protein from  
 CC PCB11. This plasmid contains the LKT 111 polypeptide fused to  
 CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity  
 CC which enables there to be an increase in antigen presentation and thus an  
 CC optimal immune response. The removal of this region also enables the  
 CC truncated LKT to be expressed at much higher levels and allows the amount  
 CC of antigen administered to be reduced. This chimeric protein comprises a  
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
 CC chimeric protein can be used as a vaccine to help reduce the incidence of  
 CC mammary tumours in a mammalian individual.  
 SQ Sequence 544 AA;

Query Match 59.6%; Score 62; DB 19; Length 544;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRGS 11  
 |||||  
 DB 495 hwsyglrpgs 504

RESULT 15  
 W79573  
 ID W79573 standard; Protein; 695 AA.  
 XX W79573;  
 AC 24-DEC-1998 (first entry)  
 DT LKT-GnRH chimeric protein.  
 XX  
 DE Chimera; PCB122; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;  
 KW cytotoxic activity; antigen presentation; immune response; vaccine;  
 KW tumour.  
 XX  
 OS Synthetic.  
 OS  
 XX MO9806848-A1.  
 PN 19-FEB-1998.  
 PD 08-AUG-1997; 97WO-CA00559.  
 PE 09-AUG-1996; 96US-0694865.  
 PR (UYSA-) UNIV SASKATCHEWAN.  
 PA Manns JG, Potter AA;  
 PI  
 XX

DR WPI; 1998-159540/14.  
 DR N-PSDB; V61535  
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -  
 PT useful for, e.g. preparation of vaccines for reduction of incidence  
 of mammary tumours in mammals  
 PS Claim 9; Figure 9.1-6; 118pp; English.  
 XX The present sequence represents the LKT-GnRH chimeric protein from  
 CC PCB122. This plasmid contains the LKT 111 polypeptide fused to sixteen  
 CC copies of the GnRH peptide. In the pattern of: 8 copies of GnRH-LKT 111-8  
 CC copies of GnRH. This chimera lacks cytotoxic activity which enables  
 CC there to be an increase in antigen presentation and thus an optimal  
 CC immune response. The removal of this region also enables the truncated  
 CC LKT to be expressed at much higher levels and allows the amount of  
 CC antigen administered to be reduced. This chimeric protein comprises a  
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
 CC chimeric protein can be used as a vaccine to help reduce the incidence of  
 CC mammary tumours in a mammalian individual.  
 SQ Sequence 695 AA;

Query Match 59.6%; Score 62; DB 19; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRGS 11  
 |||||  
 DB 10 hwsyglrpgs 19

Search completed: March 2, 2001, 10:53:46  
 Job time: 911 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:59:03 ; Search time 201.99 Seconds  
(without alignments)  
12.920 Million cell updates/sec

Title: US-09-306-689-4

Perfect score: 104

Sequence: 1 XHMSYGLRPGSSPPPC 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 957798 segs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA:\*

1:	/cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
27:	/cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
28:	/cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	103	99.0	17	1 PCT-US95-01225-6	Sequence 6, Appl
2	103	99.0	17	US-08-478-546B-6	Sequence 6, Appl
3	103	99.0	17	US-09-305-924-4	Sequence 4, Appl
4	103	99.0	17	US-09-306-689-4	Sequence 4, Appl
5	83.5	80.3	16	1 PCT-US95-01225-2	Sequence 2, Appl

6	83.5	80.3	16	8 US-08-478-546B-2	Sequence 2, Appl
7	83.5	80.3	16	17 US-09-305-924-5	Sequence 5, Appl
8	83.5	80.3	16	17 US-09-306-689-5	Sequence 5, Appl
9	65.5	63.0	41	11 US-09-506-078-23	Sequence 23, Appl
10	65.5	63.0	41	12 US-60-120-454-23	Sequence 23, Appl
11	65.5	63.0	44	19 US-09-506-078-27	Sequence 27, Appl
12	65.5	63.0	44	22 US-60-120-454-27	Sequence 27, Appl
13	62	59.6	17	5 US-08-160-882-18	Sequence 18, Appl
14	62	59.6	18	5 US-08-160-882-28	Sequence 28, Appl
15	62	59.6	49	14 US-09-019-010-4	Sequence 4, Appl
16	62	59.6	49	16 US-09-249-447A-7	Sequence 7, Appl
17	62	59.6	49	17 US-09-305-924-11	Sequence 11, Appl
18	62	59.6	49	17 US-09-306-689-11	Sequence 11, Appl
19	62	59.6	49	17 US-09-383-912-4	Sequence 4, Appl
20	62	59.6	54	17 US-09-383-912-10	Sequence 10, Appl
21	62	59.6	65	17 US-09-305-924-13	Sequence 13, Appl
22	62	59.6	65	17 US-09-306-689-13	Sequence 13, Appl
23	62	59.6	69	17 US-09-383-912-16	Sequence 16, Appl
24	62	59.6	97	17 US-09-383-912-8	Sequence 8, Appl
25	60	57.7	37	16 US-09-252-991A-20108	Sequence 20108, A
26	59	56.7	14	5 US-08-160-882-22	Sequence 22, Appl
27	59	56.7	14	5 US-08-160-882-24	Sequence 24, Appl
28	59	56.7	14	5 US-08-160-882-26	Sequence 26, Appl
29	59	56.7	18	1 PCT-US99-11219-1148	Sequence 1148, Ap
30	59	56.7	18	14 US-09-082-279-1148	Sequence 1148, Ap
31	59	56.7	18	14 US-09-082-279-1148	Sequence 1148, Ap
32	59	56.7	18	14 US-09-082-279-1148	Sequence 1148, Ap
33	59	56.7	18	17 US-09-315-304-1148	Sequence 1148, Ap
34	59	56.7	18	17 US-09-350-641-1148	Sequence 1148, Ap
35	59	56.7	18	23 PCT-US01-03988-1148	Sequence 1148, Ap
36	59	56.7	18	27 US-09-515-968-1148	Sequence 43, Appl
37	59	56.7	24	5 US-08-160-882-43	Sequence 45, Appl
38	59	56.7	44	5 US-08-160-882-45	Sequence 47, Appl
39	59	56.7	84	5 US-08-160-882-47	Sequence 153, App
40	58	55.8	9	20 US-09-657-276-153	Sequence 28, Appl
41	58	55.8	10	1 PCT-US00-22121-28	Sequence 32, Appl
42	58	55.8	10	1 PCT-US00-22121-32	Sequence 1, Appl
43	58	55.8	10	1 PCT-US94-04832A-1	Sequence 4, Appl
44	58	55.8	10	1 PCT-US94-12763-4	Sequence 1, Appl
45	58	55.8	10	27 US-09-412-558-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
PCT-US95-01225-6  
; Sequence 6, Application PC/7US9501225  
; GENERAL INFORMATION:  
; APPLICANT: Grimes, Stephen  
; TITLE OF INVENTION: Immunogens Against Gonadotropin  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dimitrios T. Drivas, Esq.  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01225  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Drivas Esq., Dimitrios T.  
; REGISTRATION NUMBER: 32,218

REFERENCE/DOCKET NUMBER: 1102865-300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-819-8286  
TELEFAX: 212-354-8113  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label=pglu  
OTHER INFORMATION: /note="pyroglutamic acid (5-oxoproline)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note="immunomimic"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11..17  
OTHER INFORMATION: /note="spacer"  
PCT-US95-01225-6

Query Match 99.0%; Score 103; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 2  
US-08-478-546B-6  
; Sequence 6, Application US/08478546B  
; GENERAL INFORMATION:  
; APPLICANT: Grimes, Stephen  
; APPLICANT: Scibbenski, Robert  
; TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent  
; TITLES OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormone  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dimitrios T. Drivas, Esq.  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,546B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/188,223  
; FILING DATE: 27-JAN-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Drivas Esq., Dimitrios T.  
; REGISTRATION NUMBER: 32,218  
; REFERENCE/DOCKET NUMBER: 1102865-300  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-819-8286  
; TELEFAX: 212-354-8113  
; INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified site  
LOCATION: 1  
OTHER INFORMATION: /note=Xaa  
OTHER INFORMATION: /note="pyroglutamic acid"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..10  
OTHER INFORMATION: /note="immunomimic"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 11..17  
OTHER INFORMATION: /note="spacer"  
US-08-478-546B-6

Query Match 99.0%; Score 103; DB 8; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 3  
US-09-305-924-4  
; Sequence 4, Application US/09305924A  
; GENERAL INFORMATION:  
; APPLICANT: Jack G. Manns  
; APPLICANT: Stephen D. Acres  
; APPLICANT: Richard Harland  
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION  
; FILE REFERENCE: 9001-0048  
; CURRENT APPLICATION NUMBER: US/09/305,924A  
; EARLIER APPLICATION NUMBER: US 60/084,217  
; EARLIER FILING DATE: 1998-05-05  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 4  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: GNRH  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)  
; OTHER INFORMATION: Xaa is pyroglutamic acid  
US-09-305-924-4

Query Match 99.0%; Score 103; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 4  
US-09-306-689-4  
; Sequence 4, Application US/09306689B  
; GENERAL INFORMATION:  
; APPLICANT: Robbins, Sarah C.  
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN  
; FILE REFERENCE: 9001-0047



;; CURRENT APPLICATION NUMBER: US/09/306,689B  
;; CURRENT FILING DATE: 1999-05-06  
;; EARLIER APPLICATION NUMBER: US 60/088,024  
;; EARLIER FILING DATE: 1998-06-04  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 4  
;; LENGTH: 17  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; NAME/KEY: MOD\_RES  
;; LOCATION: (1)  
;; OTHER INFORMATION: Xaa is pyroglutamic acid  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Sequence  
;; OTHER INFORMATION: defines GNRH analogue  
;; US-09-306-689-4

Query Match 99.0%; Score 103; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 5  
PCT-US95-01225-2  
; Sequence 2, Application PC/TUS9501225  
; GENERAL INFORMATION:  
; APPLICANT: Grimes, Stephen  
; APPLICANT: Scibienski, Robert  
; TITLE OF INVENTION: Immunogens Against Gonadotropin  
; TITLE OF INVENTION: Releasing Hormone  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Dimitrios T. Drivas, Esq.  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01225  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Drivas Esq., Dimitrios T.  
; REGISTRATION NUMBER: 32,218  
; REFERENCE/DOCKET NUMBER: 1102865-300  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-819-8286  
; TELEFAX: 212-354-8113  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: YES  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..10  
; OTHER INFORMATION: /note= "immunomimic"  
; FEATURE:  
; NAME/KEY: Region

;; NAME/KEY: Region  
;; LOCATION: 11..16  
;; OTHER INFORMATION: /note= "spacer"  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 1  
;; OTHER INFORMATION: /label= pglu  
;; OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"  
;; PCT-US95-01225-2

Query Match 80.3%; Score 83.5; DB 1; Length 16;  
Best Local Similarity 87.5%; Pred. No. 0.0019;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPG-RPPPC 16

RESULT 6  
US-08-478-546B-2  
; Sequence 2, Application US/08478546B  
; GENERAL INFORMATION:  
; APPLICANT: Grimes, Stephen  
; APPLICANT: Scibienski, Robert  
; TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent  
; TITLE OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormo  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Dimitrios T. Drivas, Esq.  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,546B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/188,223  
; FILING DATE: 27-JAN-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Drivas Esq., Dimitrios T.  
; REGISTRATION NUMBER: 32,218  
; REFERENCE/DOCKET NUMBER: 1102865-300  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-819-8286  
; TELEFAX: 212-354-8113  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified site  
; LOCATION: 1  
; OTHER INFORMATION: /note= Xaa  
; OTHER INFORMATION: /note= "pyroglutamic acid"  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..10  
; OTHER INFORMATION: /note= "immunomimic"  
; FEATURE:  
; NAME/KEY: Region

LOCATION: 11..16  
OTHER INFORMATION: /note= 'spacer'  
US-08-478-546B-2

Query Match 80.3%; Score 83.5; DB 8; Length 16;  
Best Local Similarity 87.5%; Pred. No. 0.0019;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
OY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPG-RPPPC 16

RESULT 7  
US-09-305-924-5  
Sequence 5, Application US/09305524A  
GENERAL INFORMATION:  
APPLICANT: Jack G. Manns  
APPLICANT: Stephen D. Acres  
APPLICANT: Richard Harland  
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION  
FILE REFERENCE: 9001-0048  
CURRENT APPLICATION NUMBER: US/09/305,924A  
CURRENT FILING DATE: 1999-05-05  
EARLIER APPLICATION NUMBER: US 60/084,217  
EARLIER FILING DATE: 1998-05-05  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 16  
TYPE: PRT  
ORGANISM: GNHR  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa is pyroglutamic acid  
US-09-305-924-5

Query Match 80.3%; Score 83.5; DB 17; Length 16;  
Best Local Similarity 87.5%; Pred. No. 0.0019;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
OY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPG-RPPPC 16

RESULT 8  
US-09-306-689-5  
Sequence 5, Application US/09306689B  
GENERAL INFORMATION:  
APPLICANT: Robbins, Sarah C.  
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN  
FILE REFERENCE: 9001-0047  
CURRENT APPLICATION NUMBER: US/09/306,689B  
CURRENT FILING DATE: 1999-05-06  
EARLIER APPLICATION NUMBER: US 60/088,024  
EARLIER FILING DATE: 1998-06-04  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa is pyroglutamic acid  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Sequence

OTHER INFORMATION: defines GNHR analogue  
US-09-306-689-5

Query Match 80.3%; Score 83.5; DB 17; Length 16;  
Best Local Similarity 87.5%; Pred. No. 0.0019;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
OY 2 HWSYGLRPGSSPPPC 17  
DB 2 HWSYGLRPG-RPPPC 16

RESULT 9  
US-09-506-078-23  
Sequence 23, Application US/09506078  
GENERAL INFORMATION:  
APPLICANT: Pfizer Products Inc  
TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A  
FILE REFERENCE: PCI0202A  
CURRENT APPLICATION NUMBER: US/09/506,078  
CURRENT FILING DATE: 2000-02-16  
EARLIER APPLICATION NUMBER: N/A  
EARLIER FILING DATE: 1999-02-17  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 23  
LENGTH: 411  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: 4GnRH-tm9D  
OTHER INFORMATION: encoded by pQE-GnRH:gd.  
US-09-506-078-23

Query Match 63.0%; Score 65.5; DB 19; Length 411;  
Best Local Similarity 81.2%; Pred. No. 4.5;  
Matches 13; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
OY 2 HWSYGLRPGS-SPPP 16  
DB 50 HWSYGLRPGSMSPPT 65

RESULT 10  
US-60-120-454-23  
Sequence 23, Application US/60120454  
GENERAL INFORMATION:  
APPLICANT: Campos, Manuel  
APPLICANT: Yule, Terecila D  
APPLICANT: Martind, Serge  
APPLICANT: Ditschil, Becky A  
TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A  
FILE REFERENCE: PCI0202  
CURRENT APPLICATION NUMBER: US/60/120,454  
CURRENT FILING DATE: 1999-02-17  
EARLIER APPLICATION NUMBER: N/A  
EARLIER FILING DATE: 1999-02-17  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.0 - beta  
SEQ ID NO 23  
LENGTH: 411  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: 4GnRH-tm9D  
OTHER INFORMATION: encoded by pQE-GnRH:gd.  
US-60-120-454-23

Query Match 63.0%; Score 65.5; DB 22; Length 411;  
Best Local Similarity 81.2%; Pred. No. 4.5;  
Matches 13; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 HMSGRLPGS-SPPP 16  
|||||  
DB 50 HMSGRLPGS-SLPTP 65

RESULT 11  
US-09-506-078-27  
; Sequence 27, Application US/09506078  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Products Inc  
; TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A  
; FILE REFERENCE: PC10202A  
; CURRENT APPLICATION NUMBER: US/09/506,078  
; CURRENT FILING DATE: 2000-02-16  
; EARLIER APPLICATION NUMBER: N/A  
; EARLIER FILING DATE: 1999-02-17  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 442  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
US-09-506-078-27

Query Match 63.0%; Score 65.5; DB 19; Length 442;  
Best Local Similarity 81.2%; Pred. No. 4.8;  
Matches 13; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 HMSGRLPGS-SPPP 16  
|||||  
DB 50 HMSGRLPGS-SLPTP 65

RESULT 12  
US-60-120-454-27  
; Sequence 27, Application US/60120454  
; GENERAL INFORMATION:  
; APPLICANT: Campos, Manuel  
; APPLICANT: Yule, Terecila D  
; APPLICANT: Martind, Serge  
; APPLICANT: Dirtschl, Becky A  
; TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A  
; FILE REFERENCE: PC10202  
; CURRENT APPLICATION NUMBER: US/60/120,454  
; CURRENT FILING DATE: 1999-02-17  
; EARLIER APPLICATION NUMBER: N/A  
; EARLIER FILING DATE: 1999-02-17  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.0 - beta  
; SEQ ID NO 27  
; LENGTH: 442  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: 4GnRH-tmpd-4GnRH encoded by pOE-GnRH:gd:GnRH  
US-60-120-454-27

Query Match 63.0%; Score 65.5; DB 22; Length 442;  
Best Local Similarity 81.2%; Pred. No. 4.8;  
Matches 13; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 HMSGRLPGS-SPPP 16  
|||||  
DB 50 HMSGRLPGS-SLPTP 65

RESULT 13  
US-08-160-882-18  
; Sequence 18, Application US/08160882  
; GENERAL INFORMATION:  
; APPLICANT: Russel-Jones, Gregory J.  
; APPLICANT: Stewart, Andrew G.  
; APPLICANT: Tsouis, Con G.  
; TITLE OF INVENTION: Triat/Neuropeptide-Y Fusion Proteins  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/160,882  
; FILING DATE: 03-DEC-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/690,983  
; FILING DATE: 25-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/AU90/00373  
; FILING DATE: 24-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU P3 5979  
; FILING DATE: 25-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 60042/119/BIAU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-160-882-18

Query Match 59.6%; Score 62; DB 5; Length 17;  
Best Local Similarity 90.9%; Pred. No. 0.7;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HMSGRLPGS 12  
|||||  
DB 5 HMSGRLPGS 15

RESULT 14  
US-08-160-882-28  
; Sequence 28, Application US/08160882  
; GENERAL INFORMATION:  
; APPLICANT: Russel-Jones, Gregory J.  
; APPLICANT: Stewart, Andrew G.  
; APPLICANT: Tsouis, Con G.  
; TITLE OF INVENTION: Triat/Neuropeptide-Y Fusion Proteins  
; NUMBER OF SEQUENCES: 55

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Foley & Lardner  
;; STREET: 3000 K Street, N.W., Suite 500  
;; CITY: Washington, D.C.  
;; COUNTRY: USA  
;; ZIP: 20007-5109  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/160,882  
;; FILING DATE: 03-DEC-1993  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/690,983  
;; FILING DATE: 25-JUN-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/AU90/00373  
;; FILING DATE: 24-AUG-1990  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: AU PJ 5979  
;; FILING DATE: 25-AUG-1989  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BENT, Stephen A.  
;; REGISTRATION NUMBER: 29,768  
;; REFERENCE/DOCKET NUMBER: 60042/119/BIU  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202)672-5300  
;; TELEFAX: (202)672-5399  
;; TELEX: 904136  
;;  
;; INFORMATION FOR SEQ ID NO: 28:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
US-08-160-882-28

Query Match 59.6%; Score 62; DB 5; Length 18;  
Best Local Similarity 90.9%; Pred. NO. 0.73;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSS 12  
| | | | | | | | | |  
Db 6 HWSYGLRPGSS 16

RESULT 15  
US-09-019-010-4  
;; Sequence 4, Application US/09019010  
;; GENERAL INFORMATION:  
;; APPLICANT: HARLAND, RICHARD  
;; APPLICANT: MANN, JOHN G.  
;; APPLICANT: ACRES, STEPHEN D.  
;; TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS  
;; TITLE OF INVENTION: MOLECULES  
;; NUMBER OF SEQUENCES: 6  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: ROHINS & ASSOCIATES  
;; STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
;; CITY: MENLO PARK  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94025  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/019,010  
;; FILING DATE: 05-FEB-1998  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/036,883  
;; FILING DATE: 05-FEB-1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: MCCracken, THOMAS P.  
;; REGISTRATION NUMBER: 38,548  
;; REFERENCE/DOCKET NUMBER: 9001-0035  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (650) 325-7812  
;; TELEFAX: (650) 325-7823  
;;  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 49 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
US-09-019-010-4

Query Match 59.6%; Score 62; DB 14; Length 49;  
Best Local Similarity 100.0%; Pred. NO. 1.8;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSS 11  
| | | | | | | | | |  
Db 2 HWSYGLRPGSS 11

Search completed: March 2, 2001, 10:59:03  
Job time: 373 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:59:03 ; Search time 201.99 Seconds

(without alignments)  
12.160 Million cell updates/sec

Title: US-09-306-689-5

Perfect score: 101

Sequence: 1 XHMSYGLRPGRRPPPC 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 957798 seqs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA:\*

1:	/cgn2_6/ptodata/2/paa/PCRUS.COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081.COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082.COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083.COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084.COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085.COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086.COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US099.COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US100.COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US101.COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US102.COMB.pep.*
27:	/cgn2_6/ptodata/2/paa/US103.COMB.pep.*
28:	/cgn2_6/ptodata/2/paa/US104.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	99.0	16	1	PCT-US95-01225-2
2	100	99.0	16	8	US-08-478-546B-2
3	100	99.0	16	17	US-09-305-924-5
4	100	99.0	16	17	US-09-306-689-5
5	83.5	82.7	17	1	PCT-US95-01225-6

6	83.5	82.7	17	8	US-08-478-546B-6	Sequence 6, Appli
7	83.5	82.7	17	17	US-09-305-924-4	Sequence 4, Appli
8	83.5	82.7	17	17	US-09-306-689-4	Sequence 4, Appli
9	60	59.4	16	17	US-09-341-590-103	Sequence 103, App
10	60	59.4	16	17	US-09-341-590-104	Sequence 104, App
11	59	58.4	20	1	PCT-US00-22121-26	Sequence 26, Appl
12	59	58.4	20	3	US-07-761-849-10	Sequence 10, Appl
13	59	58.4	20	8	US-08-476-013-10	Sequence 10, Appl
14	59	58.4	20	8	US-08-476-013-12	Sequence 12, Appl
15	59	58.4	20	8	US-08-477-298-10	Sequence 10, Appl
16	59	58.4	20	14	US-09-026-276-26	Sequence 26, Appl
17	59	58.4	20	14	US-09-026-276-30	Sequence 30, Appl
18	59	58.4	20	16	US-09-214-009-1	Sequence 1, Appli
19	59	58.4	20	17	US-09-374-721A-26	Sequence 26, Appl
20	59	58.4	20	17	US-09-374-721A-30	Sequence 30, Appl
21	59	58.4	21	3	US-07-761-849-11	Sequence 11, Appl
22	59	58.4	21	3	US-07-761-849-12	Sequence 12, Appl
23	59	58.4	21	8	US-08-476-013-11	Sequence 11, Appl
24	59	58.4	21	8	US-08-476-013-12	Sequence 12, Appl
25	59	58.4	21	8	US-08-476-013-13	Sequence 13, Appl
26	59	58.4	21	8	US-08-476-013-14	Sequence 14, Appl
27	59	58.4	21	8	US-08-477-298-11	Sequence 11, Appl
28	59	58.4	21	8	US-08-477-298-14	Sequence 14, Appl
29	59	58.4	40	1	PCT-US00-22121-35	Sequence 35, Appl
30	59	58.4	40	14	US-09-026-276-35	Sequence 35, Appl
31	59	58.4	40	17	US-09-374-721A-35	Sequence 35, Appl
32	59	58.4	41	1	PCT-US00-22121-34	Sequence 34, Appl
33	59	58.4	41	14	US-09-026-276-34	Sequence 34, Appl
34	59	58.4	41	19	US-09-374-721A-34	Sequence 23, Appl
35	59	57.9	411	22	US-60-120-454-23	Sequence 23, Appl
36	58.5	57.9	442	19	US-09-506-078-27	Sequence 27, Appl
37	58.5	57.9	442	20	US-09-506-078-27	Sequence 27, Appl
38	58.5	57.9	442	20	US-09-506-078-27	Sequence 27, Appl
39	58.5	57.9	442	20	US-09-506-078-27	Sequence 27, Appl
40	58	57.4	9	22	US-09-657-276-153	Sequence 153, Appl
41	58	57.4	10	1	PCT-US00-22121-32	Sequence 32, Appl
42	58	57.4	10	1	PCT-US00-22121-32	Sequence 32, Appl
43	58	57.4	10	1	PCT-US94-04832A-1	Sequence 1, Appli
44	58	57.4	10	1	PCT-US94-12763-4	Sequence 1, Appli
45	58	57.4	10	27	US-09-412-558-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
PCT-US95-01225-2  
Sequence 2, Application PC/FUS9501225  
GENERAL INFORMATION:  
APPLICANT: Grimes, Stephen  
APPLICANT: Sciablenki, Robert  
TITLE OF INVENTION: Immunogens Against Gonadotropin  
TITLE OF INVENTION: Releasing Hormone  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dimitrios T. Drivas, Esq.  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01225  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Drivas Esq., Dimitrios T.  
REGISTRATION NUMBER: 32,218

```
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..16
OTHER INFORMATION: /note="spacer"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label="cglu"
OTHER INFORMATION: /note="pyroglutamic acid (5-oxoproline)"
PCT-US95-01225-2
```

```
Query Match          99.0%; Score 100; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HWSYGLRPGRRPPPC 16
Db      2 HWSYGLRPGRRPPPC 16
```

```
RESULT 2
US-08-478-546B-2
; Sequence 2, Application US/08478546B
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scibienski, Robert
; TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent
; TITILE OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,546B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/188,223
; FILING DATE: 27-JAN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286
; TELEFAX: 212-354-8113
; INFORMATION FOR SEQ ID NO: 2:
```

```
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified site
LOCATION: 1
OTHER INFORMATION: /note="xaa"
OTHER INFORMATION: /note="pyroglutamic acid"
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..16
OTHER INFORMATION: /note="spacer"
US-08-478-546B-2
```

```
Query Match          99.0%; Score 100; DB 8; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HWSYGLRPGRRPPPC 16
Db      2 HWSYGLRPGRRPPPC 16
```

```
RESULT 3
US-09-305-924-5
; Sequence 5, Application US/09305924A
; GENERAL INFORMATION:
; APPLICANT: Stephen D. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; EARLIER FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: GARRH
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)
; OTHER INFORMATION: Xaa is pyroglutamic acid
US-09-305-924-5
```

```
Query Match          99.0%; Score 100; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 HWSYGLRPGRRPPPC 16
Db      2 HWSYGLRPGRRPPPC 16
```

```
RESULT 4
US-09-306-689-5
; Sequence 5, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; ANIMALS
; FILE REFERENCE: 9001-0047
```

;;  
; CURRENT APPLICATION NUMBER: US/09/306,689B  
; CURRENT FILING DATE: 1999-05-06  
; EARLIER APPLICATION NUMBER: US 60/088,024  
; EARLIER FILING DATE: 1998-06-04  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)  
; OTHER INFORMATION: Xaa is pyroglutamic acid  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Sequence  
; OTHER INFORMATION: defines GNRH analogue  
; US-09-306-689-5

Query Match 99.0%; Score 100; DB 17; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRRPPPC 16  
|||||  
DB 2 HWSYGLRPGRRPPPC 16

RESULT 5  
PCT-US95-01225-6  
; Sequence 6, Application PC/TUS9501225  
; GENERAL INFORMATION:  
; APPLICANT: Grimes, Stephen  
; APPLICANT: Scibienski, Robert  
; TITLE OF INVENTION: Immunogens Against Gonadotropin  
; TITLE OF INVENTION: Releasing Hormone  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dimitrios T. Drivas, Esq.  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01225  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Drivas Esq., Dimitrios T.  
; REGISTRATION NUMBER: 32,218  
; REFERENCE/DOCKET NUMBER: 1102865-300  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-819-8286  
; TELEFAX: 212-354-8113  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: YES  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1  
; OTHER INFORMATION: /label="pglu  
; OTHER INFORMATION: /note="pyroglutamic acid (5-oxoproline)"

;;  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..10  
; OTHER INFORMATION: /note="immunomimic"  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 11..17  
; OTHER INFORMATION: /note="spacer"  
; PCT-US95-01225-6

Query Match 82.7%; Score 83.5; DB 1; Length 17;  
Best Local Similarity 87.5%; Pred. No. 0.0017;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 2 HWSYGLRPG-RPPPC 16  
|||||  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 6  
US-08-478-546B-6  
; Sequence 6, Application US/08478546B  
; GENERAL INFORMATION:  
; APPLICANT: Grimes, Stephen  
; APPLICANT: Scibienski, Robert  
; TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent  
; TITLE OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormo  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dimitrios T. Drivas, Esq.  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,546B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/188,223  
; FILING DATE: 27-JAN-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Drivas Esq., Dimitrios T.  
; REGISTRATION NUMBER: 32,218  
; REFERENCE/DOCKET NUMBER: 1102865-300  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-819-8286  
; TELEFAX: 212-354-8113  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified site  
; LOCATION: 1  
; OTHER INFORMATION: /note="Xaa  
; OTHER INFORMATION: /note="pyroglutamic acid"  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..10  
; OTHER INFORMATION: /note="immunomimic"  
; FEATURE:  
; NAME/KEY: Region

LOCATION: 11..17  
OTHER INFORMATION: /note="spacer"  
US-08-478-546B-6

Query Match 82.7%; Score 83.5; DB 8; Length 17;  
Best Local Similarity 87.5%; Pred. No. 0.0017;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 2 HWSYGLRPG-RPPPC 16  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 7  
US-09-305-924-4  
Sequence 4, Application US/09305924A  
GENERAL INFORMATION:  
APPLICANT: Jack G. Manns  
APPLICANT: Stephen D. Acres  
APPLICANT: Richard Harland  
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION  
FILE REFERENCE: 9001-0048  
CURRENT APPLICATION NUMBER: US/09/305,924A  
CURRENT FILING DATE: 1999-05-05  
EARLIER APPLICATION NUMBER: US 60/084,217  
EARLIER FILING DATE: 1998-05-05  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 4  
LENGTH: 17  
TYPE: PRT  
ORGANISM: GNRH  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa is pyroglutamic acid  
US-09-305-924-4

Query Match 82.7%; Score 83.5; DB 17; Length 17;  
Best Local Similarity 87.5%; Pred. No. 0.0017;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 2 HWSYGLRPG-RPPPC 16  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 8  
US-09-306-689-4  
Sequence 4, Application US/09306689B  
GENERAL INFORMATION:  
APPLICANT: Robbins, Sarah C.  
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN  
TITLE OF INVENTION: ANIMALS  
FILE REFERENCE: 9001-0047  
CURRENT APPLICATION NUMBER: US/09/306,689B  
CURRENT FILING DATE: 1999-05-06  
EARLIER APPLICATION NUMBER: US 60/088,024  
EARLIER FILING DATE: 1998-06-04  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 4  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa is pyroglutamic acid  
OTHER INFORMATION: Description of Artificial Sequence: Sequence

OTHER INFORMATION: defines GNRH analogue  
US-09-306-689-4

Query Match 82.7%; Score 83.5; DB 17; Length 17;  
Best Local Similarity 87.5%; Pred. No. 0.0017;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 2 HWSYGLRPG-RPPPC 16  
DB 2 HWSYGLRPGSSPPPC 17

RESULT 9  
US-09-341-590-103  
Sequence 103, Application US/09341590  
GENERAL INFORMATION:  
APPLICANT: Larsen, Bjarne Due  
TITLE OF INVENTION: PHARMACOLOGICALLY ACTIVE PEPTIDE CONJUGATES HAVING A  
TITLE OF INVENTION: REDUCED TENDENCY TOWARDS ENZYMATIC HYDROLYSIS  
FILE REFERENCE: PPT-20479-US  
CURRENT APPLICATION NUMBER: US/09/341,590  
CURRENT FILING DATE: 1999-07-03  
PRIOR APPLICATION NUMBER: DK 0317/98  
PRIOR FILING DATE: 1998-03-09  
NUMBER OF SEQ ID NOS: 121  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 103  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: GNRH-Lys6  
US-09-341-590-103

Query Match 59.4%; Score 60; DB 17; Length 16;  
Best Local Similarity 90.0%; Pred. No. 0.98;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGR 11  
DB 2 HWSYGLRPGR 11

RESULT 10  
US-09-341-590-104  
Sequence 104, Application US/09341590  
GENERAL INFORMATION:  
APPLICANT: Larsen, Bjarne Due  
TITLE OF INVENTION: PHARMACOLOGICALLY ACTIVE PEPTIDE CONJUGATES HAVING A  
TITLE OF INVENTION: REDUCED TENDENCY TOWARDS ENZYMATIC HYDROLYSIS  
FILE REFERENCE: PPT-20479-US  
CURRENT APPLICATION NUMBER: US/09/341,590  
CURRENT FILING DATE: 1999-07-03  
PRIOR APPLICATION NUMBER: DK 0317/98  
PRIOR FILING DATE: 1998-03-09  
NUMBER OF SEQ ID NOS: 121  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 104  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: GNRH-(Lys-Gly)<sup>3</sup>  
US-09-341-590-104

Query Match 59.4%; Score 60; DB 17; Length 16;  
Best Local Similarity 90.0%; Pred. No. 0.98;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 HWSYGLRPGR 11



Db 2 HWSYGLRPGQ 11

RESULT 11  
PCT-US00-22121-26  
; Sequence 26, Application PC/TUS0022121  
; GENERAL INFORMATION:  
; APPLICANT: Proteinix Company  
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM  
; FILE REFERENCE: IGN-2004WO  
; CURRENT APPLICATION NUMBER: PCT/US00/22121  
; CURRENT FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 09/026,276  
; PRIOR FILING DATE: 1998-02-19  
; PRIOR APPLICATION NUMBER: US 09/374,721  
; PRIOR FILING DATE: 1999-08-13  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide  
PCT-US00-22121-26

Query Match 58.4%; Score 59; DB 1; Length 20;  
Best Local Similarity 90.0%; Pred. No. 1.6;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 2 HWSYGLRPGR 11  
Db 2 HWSYGLRPGQ 11

RESULT 12  
PCT-US00-22121-30  
; Sequence 30, Application PC/TUS0022121  
; GENERAL INFORMATION:  
; APPLICANT: Proteinix Company  
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM  
; FILE REFERENCE: IGN-2004WO  
; CURRENT APPLICATION NUMBER: PCT/US00/22121  
; CURRENT FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 09/026,276  
; PRIOR FILING DATE: 1998-02-19  
; PRIOR APPLICATION NUMBER: US 09/374,721  
; PRIOR FILING DATE: 1999-08-13  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 30  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide  
PCT-US00-22121-30

Query Match 58.4%; Score 59; DB 1; Length 20;  
Best Local Similarity 90.0%; Pred. No. 1.6;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 2 HWSYGLRPGR 11  
Db 2 HWSYGLRPGQ 11  
RESULT 13

US-07-761-849-10  
; Sequence 10, Application US/07761849  
; GENERAL INFORMATION:  
; APPLICANT: Meleoen, Robert H.  
; TITLE OF INVENTION: WENSING, CORNELIUS J. G.  
; TITLE OF INVENTION: PEPTIDE, IMMUNOGENIC COMPOSITION AND VACCINE  
; TITLE OF INVENTION: OR MEDICINAL PREPARATION: A METHOD OF  
; TITLE OF INVENTION: IMMUNISING A MAMMAL AGAINST LHRH, AND A METHOD  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/761,849  
; FILING DATE: 19910917  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Moran, Thomas F.  
; REGISTRATION NUMBER: 16,579  
; REFERENCE/DOCKET NUMBER: 2805/40057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-07-761-849-10

Query Match 58.4%; Score 59; DB 3; Length 20;  
Best Local Similarity 90.0%; Pred. No. 1.6;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 2 HWSYGLRPGR 11  
Db 2 HWSYGLRPGQ 11

RESULT 14  
US-08-476-013-10  
; Sequence 10, Application US/08476013  
; GENERAL INFORMATION:  
; APPLICANT: Meleoen, Robert H.  
; TITLE OF INVENTION: WENSING, CORNELIUS J. G.  
; TITLE OF INVENTION: PEPTIDE, IMMUNOGENIC COMPOSITION AND VACCINE  
; TITLE OF INVENTION: OR MEDICINAL PREPARATION: A METHOD OF  
; TITLE OF INVENTION: IMMUNISING A MAMMAL AGAINST LHRH, AND A METHOD  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,013  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Katz, Robert D.  
REGISTRATION NUMBER: 30,141  
REFERENCE/DOCKET NUMBER: 40057-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-476-013-10

Query Match 58.4%; Score 59; DB 8; Length 20;  
Best Local Similarity 90.0%; Pred. No. 1.6;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 HWSYGLRPGR 11  
Db 2 HWSYGLRPGQ 11

## RESULT 15

US-08-476-013-12

Sequence 12, Application US/08476013A

GENERAL INFORMATION:

APPLICANT: Meislen, Robert H

APPLICANT: Mensing, Cornelius J. G.

TITLE OF INVENTION: A PEPTIDE, IMMUNOGENIC COMPOSITIONS, VACCINES, AND

TITLE OF INVENTION: MEDICINAL PREPARATIONS INCLUDING SAME; AND METHODS

TITLE OF INVENTION: FOR IMMUNIZING A MAMMAL AGAINST LHRH, AND IMPROVING

TITLE OF INVENTION: THE MEAT QUALITY OF PIGS

FILE REFERENCE: 2185-403605

CURRENT APPLICATION NUMBER: US/08/476,013A

CURRENT FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 12

LENGTH: 20

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: at least two

OTHER INFORMATION: LHRH peptide sequences in tandem

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)

OTHER INFORMATION: Residue may be Pyroglutamic Acid or Gln having

OTHER INFORMATION: attached thereto a tail including one or more

OTHER INFORMATION: additional amino acids

FEATURE:

NAME/KEY: VARIANT

LOCATION: (3)

OTHER INFORMATION: Trp may be formylated

FEATURE:

NAME/KEY: VARIANT

LOCATION: (13)

OTHER INFORMATION: TRP may be formylated

FEATURE:

NAME/KEY: VARIANT

LOCATION: (10) (19)

OTHER INFORMATION: Sequence may repeat creating more than two LHRH

OTHER INFORMATION: sequences in tandem

FEATURE:

NAME/KEY: VARIANT  
LOCATION: (10) (11)  
OTHER INFORMATION: Gly residue may be joined to Gln residue by a  
OTHER INFORMATION: direct bond or by a spacer made of one or more  
OTHER INFORMATION: amino acids, a hydrocarbon chain or other groups  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (20)  
OTHER INFORMATION: Residue may be amidated or may have attached  
OTHER INFORMATION: thereto a tail of one or more additional amino  
OTHER INFORMATION: acids  
US-08-476-013-12

Query Match 58.4%; Score 59; DB 8; Length 20;  
Best Local Similarity 90.0%; Pred. No. 1.6;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 HWSYGLRPGR 11  
Db 2 HWSYGLRPGQ 11

Search completed: March 2, 2001, 10:59:03  
Job time: 373 sec



PR 27-JAN-1994; 94US-0188223.  
XX  
PA (APH-T-) APHTON CORP.  
XX  
PI Grimes S, Scibienski R.  
XX  
DR WPI; 1995-275410/36.  
XX  
PT New anti-gonadotropin releasing hormone immunogenic composition(s)  
PT used for treating gonadotropin and gonadal steroid hormone  
PT dependent disease(s) and providing contraception.  
XX  
PS  
XX  
XX Claim 1; Page 29; 39pp; English.  
XX  
XX The sequences given in R78282-85 represent immunoinimetics to  
CC gonadotropin releasing hormone (GnRH). These peptides comprise the  
CC wild type GnRH sequence and a spacer attached to either the N- or C-  
CC terminal. The spacer molecules serve as a link through which the  
CC immunoinimic is attached to an immunological carrier such as  
CC diphtheria toxoid (DT) and also affects the immune response generated  
CC by the vaccinated mammal against the immunoinimic. Compositions  
CC comprising these peptides may be used for treating a mammal for  
CC gonadotropin and gonadal steroid hormone associated dependent  
CC disease or for providing immunological contraception in mammals.  
CC CC This can also be used for treating breast cancer, uterine and other  
CC gynaecological cancers, endometriosis, uterine fibroids, prostate  
CC cancer, or benign prostatic hypertrophy.  
XX  
XX  
SQ Sequence 16 AA;

Query Match	99.0%	Score 100:	DB 16:	Length 16:
Best Local Similarity	100.0%	Fred. No.	7.2e-07:	
Matches 15, Conservative	0:	Mismatches	0:	Indels 0:
				Gaps 0:

QY	2	HWSYGLRPGRRPPPC	16
Db	2	hwsyglrpgrrpppc	16

RESULT	2
Y58141	
ID	Y58141 standard; peptide; 16 AA.

AC Y58141;  
XX  
DT 07-MAR-2000 (first entry)

**KW** Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein,  
soluble form of gp120  
**DE** Gonadotropin releasing hormone (GnRh) peptide analogue 5.  
**XX**

KW non-androgenic; steroid; reduction; weight gain; muscle distribution;  
KW fat distribution; male pattern; boar taint; flavour; impairment;  
KW . reliable; immunocastration; meat production.

	key	Location/Qualifiers
US	synthetic.	
XX		
FH		
ET	Modified-site	1

	fylogiuramic acid
FI	
XX	
PN	
XX	
	W09956771-A2.

PD	11-NOV-1999.
XX	
XX	05-MAY-1999;
PF	99WO-CA00360.
XX	

PA (BIOS-) BIOSTAR INC.  
XX  
XX

PI Manns JG, Acres SD, Harland F;

XX WPI; 2000-062125/05.  
XX  
XX  
XX Production of uncastrated male food animals using vaccines -  
XX  
XX  
XX Disclosure; Page 11; 87pp; English.  
PS

CC SequencingY58136-Y58141 represent gonadotropin releasing hormone  
CC (GnRH) analogues which may be used as an alternative to sequence  
CC Y58133 in embodiments of the present invention. The invention  
CC relates to a method of using two GnRH immunogen vaccines to produce  
CC uncastrated male animals for meat production, one vaccination prior to  
CC or during the fattening period to reduce circulating testosterone levels  
CC and the second vaccination about 2-8 weeks before slaughter to  
CC substantially reduce androgenic and/or non-androgenic steroids. The  
CC invention is used to produce food animals that exhibit the weight gain  
CC and muscle/fat distribution of male animals without the problems  
CC associated with male animals. Such problems include "boar taint", a  
CC urine-like odour found in cooked meat of uncastrated pigs which is  
CC caused by steroids stored in the tissues, and similar flavour  
CC impairments in the meat of other intact male animals. The invention is  
CC more reliable than prior art immunocastration techniques.

Sequence	16 AA
<p>           SQ            Sequence         </p>	<p>           16 AA         </p>

Query Match	99.0%	Score 100:	DB 21:	Length 16:
Best Local Similarity	100.0%	Pred. No.	7.2e-07:	
Matches 15, Conservative	0:	Mismatches	0:	Indels 0
				Caps 0

QY	2	HWSYGLRPGRRPPPC	16
Db	2	hwsyglrpgrrppppc	16

RESULT	3
R78283	
ID	R78283 standard; peptide; 17 AA

AC	R78283;
XX	
DT	13-MAR-1996 (first entry)

De GnRH immunomimetic and spacer (GnRH(1-10)-Ser10)  
XX  
KW Immunomimetic; gonadotropin releasing hormone;  
Intestine; hormone release; hormone; GnRH

gonadal steroid hormone associated dependent disease; gonadotropin;  
immunological contraception; mammal; breast; cancer; prostate;  
hermaphroditic; hermaphrodite;

XX	chimeric - Homo sapiens.
OS	
OS	chimeric - Synthetic.
XY	

FT	Modified-site	1	location/qualifiers
FT	key	OTHER	/note= "nvroq]utami
FT			

FT	Peptide	"note="	"spacer"
FT	FT	1.10	"GNRH"
FT	FT	11.17	"note="

AA W09520600-A1.  
PN  
XX  
PD 03-AUG-1995.

AA		
PF	26-JAN-1995;	95WO-US01225.
XX		
PR	27-JAN-1994;	94US-0188223.

PA (ARHT-) ARHTON CORP.

PI	Grimes S, Sciablenki R:
XX	WPI: 1995-275410/36.
XX	
PT	New anti-gonadotropin releasing hormone immunogenic composition(s)
PT	- used for treating gonadotropin and gonadal steroid hormone
PT	dependent disease(s) and providing contraception.
XX	
PS	Claim 1; Page 29; 39pp; English.
XX	
CC	The sequences given in R78282-85 represent immunomimetics to
CC	gonadotropin releasing hormone (GnRH). These peptides comprise the
CC	wild type GnRH sequence and a spacer attached to either the N- or C-
CC	terminal. The spacer molecules serve as a link through which the
CC	immunomimic is attached to an immunological carrier such as
CC	diphtheria toxoid (DT) and also affects the immune response generated
CC	by the vaccinated mammal against the immunomimic. Compositions
CC	comprising these peptides may be used for treating a mammal for
CC	gonadotropin and gonadal steroid hormone associated dependent
CC	disease or for providing immunological contraception in mammals.
CC	They can also be used for treating breast cancer, uterine and other
CC	gynaecological cancers, endometriosis, uterine fibroids, prostate
CC	cancer, or benign prostatic hypertrophy.
XX	
SO	Sequence 17 AA:
	Query Match 82.7%; Score 83.5; DB 16; Length 17;
	Best Local Similarity 87.5%; Pred. No. 9.3e-05;
	Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1
OY	2 HWSYGLRPG-RPPPC 16
	11111
Db	2 hwsyglrpgssp PPC 17
RESULT 4	
ID	Y58140 standard; peptide; 17 AA.
XX	
AC	Y58140;
XX	
DT	07-MAR-2000 (first entry)
XX	
DE	Gonadotropin releasing hormone (GnRH) peptide analogue 4.
XX	
KW	Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fission protein;
KW	antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
KW	non-androgenic; steroid; reduction; weight gain; muscle distribution;
KW	fat distribution; male pattern; boar taint; flavour; impairment;
XX	reliable; immunocastration; meat production.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	1
FT	Modified-site /note= "Pyroglutamic acid"
XX	
PN	W09956771-A2.
PD	
PD	11-NOV-1999.
XX	
PF	05-MAY-1999; 99WO-CA00360.
XX	
PR	05-MAY-1998; 98US-0084217.
XX	
PA	(BIOS-) BIOSTAR INC.
XX	
PI	Manns JG, Acres SD, Harland R;
XX	
DR	WPI: 2000-062125/05.
XX	

PT Production of uncastrated male food animals using vaccines -  
XX  
PS Disclosure; Page 11; 87pp; English.  
XX  
CC Sequences Y58136-Y58141 represent gonadotropin releasing hormone  
CC (GnRH) analogues which may be used as an alternative to sequence  
CC Y58135 in embodiments of the present invention. The invention  
CC relates to a method of using two GnRH immunogen vaccines to produce  
CC uncastrated male animals for meat production, one vaccination prior to  
CC or during the fattening period to reduce circulating testosterone levels,  
CC and the second vaccination about 2-8 weeks before slaughter to  
CC substantially reduce androgenic and/or non-androgenic steroids. The  
CC invention is used to produce food animals that exhibit the weight gain  
CC and muscle/fat distribution of male animals without the problems  
CC associated with male animals. Such problems include "boar taint", a  
CC urine-like odour found in cooked meat of uncastrated pigs which is  
CC caused by steroids stored in the tissues, and similar flavour  
CC impairments in the meat of other intact male animals. The invention is  
CC more reliable than prior art immunocastration techniques.  
XX  
SQ Sequence 17 AA;  
  
Query Match 82.7%; Score 83.5; DB 21; Length 17;  
Best Local Similarity 87.5%; Pred. No. 9.3e-05;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1  
  
QY 2 HMSYGLRPG-RPPPC 16  
||||||| |||||  
Db 2 hwsyglrpgsspppc 17  
  
RESULT 5  
R44955  
ID R44955 standard; Protein; 15 AA.  
XX  
AC R44955;  
XX  
DT 07-JUL-1994 (first entry)  
XX  
DE E. coli fimbriae HR4 GnRH insert linker.  
XX  
KW Gonadotropin releasing hormone; immunogenic carrier system;  
KW immune response; immunisation; vaccine; contraceptive agent;  
KW treatment; sexual hyperactivity; cancer.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 2..11  
FT /note="decapeptide GnRH"  
XX  
PN EP578293-A.  
XX  
PD 12-JAN-1994.  
XX  
PE 15-JUN-1993; 93EP-0201712.  
XX  
PR 18-JUN-1992; 92EP-0201775.  
XX  
PA (ALKU ) AKZO NV.  
XX  
PI Gielen JT, Hoekstra WPM, Van Der Zee A, Van Die IM;  
XX  
DR WPI: 1994-010165/02.  
XX  
N-PSDB; 054394.  
XX  
XX Immunogenic carrier system comprising Escherichia coli P-fimbriae  
PT - used to elicit immune response against gonadotropin releasing  
PT hormone  
XX  
XX Example 1; Fig 2; 34pp; English.  
XX

CC The sequence is that of an insert contg. the decapeptide gonadotropin  
 CC releasing hormone (GnRH) located in the major subunit of P-fimbrinae  
 CC with serotype F11 at a position corresponding to a position in  
 CC hypervariable region 4 of the wild type major subunit. This forms  
 CC part of an immunogenic carrier system capable of eliciting an immune  
 CC response against GnRH.

SQ Sequence 15 AA;

Query Match 62.4%; Score 63; DB 15; Length 15;

Best Local Similarity 90.9%; Pred. No. 0.033; 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPRG 12  
 |||||  
 DB 4 hwsyglrpgdp 14

RESULT 6

R44956 R44956 standard; Protein; 16 AA.

AC R44956;

DT 07-JUL-1994 (first entry)

XX E. coli fimbrinae HR4 GnRH insert linker.

DE Gonadotropin releasing hormone; immunogenic carrier system;

KW Immune response; Immunisation; vaccine; contraceptive agent;

KM treatment; sexual hyperactivity; cancer.

XX Synthetic.

OS Key

FT Peptide

FT Location/Qualifiers

FT 2..11

FT /note= "decapeptide GnRH"

PN EP578293-A.

PD 12-JAN-1994.

PF 15-JUN-1993; 93EP-0201712.

PR 18-JUN-1992; 92EP-0201775.

XX (ALKU) AKZO NV.

PI Gielen JT, Hoekstra WPM, Van Der Zee A, Van Die IM;

DR WPI; 1994-010165/02.

DR N-PSDB; Q54395.

XX Immunogenic carrier system comprising Escherichia coli P-fimbrinae

PT - used to elicit immune response against gonadotropin releasing

PT hormone

XX Example 1; Fig 2; 34pp; English.

XX The sequence is that of an insert contg. the decapeptide gonadotropin

CC releasing hormone (GnRH) located in the major subunit of P-fimbrinae

CC with serotype F11 at a position corresponding to a position in

CC hypervariable region 4 of the wild type major subunit. This forms

CC part of an immunogenic carrier system capable of eliciting an immune

CC response against GnRH.

XX Sequence 16 AA;

Query Match 62.4%; Score 63; DB 15; Length 16;

Best Local Similarity 90.9%; Pred. No. 0.035; 1; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPRG 12  
 |||||  
 DB 5 hwsyglrpgdp 15

RESULT 7

W47438 W47438 standard; peptide; 20 AA.

AC W47438;

DT 05-JUN-1998 (first entry)

XX Antigenic peptide.

DE Vaccine; antigen.

XX Synthetic.

OS WO9749425-A1.

PN 31-DEC-1997.

PD 24-JUN-1997; 97WO-NL00354.

PF 25-JUN-1996; 96EP-0201766.

PR (DAVE-) DANISH VETERINARY INST ANIMAL VIRUS RES.

PA (DIER-) STICHTING INST DIERHOUDERT EN DIERGEZONDHEID.

XX Beekman NJCM, Dalsgaard K, Melsen RH, Schaaper WMM;

DR WPI; 1998-076912/07.

XX Vaccines comprising antigen bound to carrier by an in vivo labile

PT bond - especially synthetic peptide linked to fatty acid via

PT thioester or di: sulphide, provide greater immune response for weakly

PT immunogenic antigens

XX Claim 9; Page 28; 36pp; English.

PS A novel vaccine comprises an antigen (Ag), e.g. the present

XX peptide, and carrier connected by a bond that is labile and

CC dissociates under certain physiological conditions.

CC The vaccine, which allows dissociation of the Ag from the carrier

CC molecule, can be used to elicit better immune responses against

CC poorly immunogenic Ag then those Ag which contain a stable link to

CC the carrier molecule. Ag dissociate from the carrier in vivo,

CC resulting in better immune response for Ag that are normally only

CC weakly immunogenic. The vaccine also improves targeting to, and

CC presentation by Ag-presenting cells.

XX Sequence 20 AA;

Query Match 58.4%; Score 59; DB 19; Length 20;

Best Local Similarity 90.0%; Pred. No. 0.14; 1; Indels 0; Gaps 0;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPRG 11  
 |||||  
 DB 2 hwsyglrpgdp 11

RESULT 8

Y31174 Y31174 standard; peptide; 20 AA.

AC Y31174;

DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH dimer for C-terminal extension.  
 XX  
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;  
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;  
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;  
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;  
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;  
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;  
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;  
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9942472-A1.  
 XX  
 PD 26-AUG-1999.  
 XX  
 PF 26-JAN-1999; 99WO-US01588.  
 XX  
 PR 19-FEB-1998; 98US-0026276.  
 XX  
 PA (IGEN-) IGEN INT INC.  
 XX  
 PI Kenten JH, Lohas GL, Pilon AL, Roberts SF, Tramontano A;  
 XX  
 DR WPI; 1999-518582/43.  
 XX  
 PT Epitope-containing fusion proteins used to generate a highly  
 XX specific immune responses  
 XX  
 PS Example 3; Page 40; 67pp; English.  
 XX  
 CC This invention describes a novel fusion protein, comprising a heat shock  
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner  
 CC which is useful for the stimulation of a highly specific immune response  
 CC when administered to an animal. The protein of the invention may be  
 CC post-translationally modified (e.g. by the addition of fatty acids to  
 CC enhance immunogenicity). The fusion proteins of the invention can be  
 CC used as vaccines to induce an immune response. When a T cell epitope is  
 CC attached, they can be used for control of viral infections, bacterial  
 CC infections, parasitic infection and cancer. The fusion proteins can be  
 CC used in pharmaceutical compositions for the treatment of gastrointestinal  
 CC diseases, pulmonary infections, respiratory infections, and HIV  
 CC infections. The use of ubiquitin as a scaffold is also useful for the  
 CC presentation and stimulation of anti-self immune responses, e.g.  
 CC generation of anti-gonadotropin releasing hormone antibodies which result  
 CC in the suppression of luteinizing hormone and follicle stimulating  
 CC hormone. This indirectly suppresses steriodogenesis and gamete maturation  
 CC in males and females. This type of anti-self response in humans is useful  
 CC in the treatment of prostate cancer and breast cancer. In livestock, the  
 CC ability to stimulate an anti-self response provides a simple alternative  
 CC to physical castration. Immunocastration of pigs is a better alternative  
 CC to physical castration, as it does not result in any of the detrimental  
 CC side effects associated with physical castration. Other examples of  
 CC diseases and conditions treated with self proteins fused with ubiquitin  
 CC are TNF and its epitopes to modulate septic shock, arthritis,  
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig  
 CC epsilon heavy chain for the control of allergic reactions; chorionic  
 CC gonadotropin for fertility control; and sperm proteins for fertility  
 CC control. A further use of the fusion proteins is as part of a vaccine to  
 CC enhance growth rate and thereby the final weight of the livestock prior  
 CC to shipment to market. In addition, the fusion proteins of the invention  
 CC can be used to detect and identify antibodies from experimental samples.  
 CC This sequence represents a GnRH dimer used in the construction of a  
 CC ubiquitin fusion protein described in the method of the invention.  
 CC  
 XX  
 SQ Sequence 20 AA:

Query Match 58.4%; Score 59; DB 20; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 0.14;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPCR 11  
 Db 2 hwsyglrpgq 11  
 |||||  
 RESULT 9  
 Y31178  
 ID Y31178 standard; peptide: 20 AA.  
 XX  
 AC Y31178;  
 XX  
 DT 28-OCT-1999 (first entry)  
 XX  
 DE Ubiquitin fusion protein GnRH mixed dimer 1.  
 XX  
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;  
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;  
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;  
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;  
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;  
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;  
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;  
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9942472-A1.  
 XX  
 PD 26-AUG-1999.  
 XX  
 PF 26-JAN-1999; 99WO-US01588.  
 XX  
 PR 19-FEB-1998; 98US-0026276.  
 XX  
 PA (IGEN-) IGEN INT INC.  
 XX  
 PI Kenten JH, Lohas GL, Pilon AL, Roberts SF, Tramontano A;  
 XX  
 DR WPI; 1999-518582/43.  
 XX  
 PT Epitope-containing fusion proteins used to generate a highly  
 XX specific immune responses  
 XX  
 PS Example 3; Page 41; 67pp; English.  
 XX  
 CC This invention describes a novel fusion protein, comprising a heat shock  
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner  
 CC which is useful for the stimulation of a highly specific immune response  
 CC when administered to an animal. The protein of the invention may be  
 CC post-translationally modified (e.g. by the addition of fatty acids to  
 CC enhance immunogenicity). The fusion proteins of the invention can be  
 CC used as vaccines to induce an immune response. When a T cell epitope is  
 CC attached, they can be used for control of viral infections, bacterial  
 CC infections, parasitic infection and cancer. The fusion proteins can be  
 CC used in pharmaceutical compositions for the treatment of gastrointestinal  
 CC diseases, pulmonary infections, respiratory infections, and HIV  
 CC infections. The use of ubiquitin as a scaffold is also useful for the  
 CC presentation and stimulation of anti-self immune responses, e.g.  
 CC generation of anti-gonadotropin releasing hormone antibodies which result  
 CC in the suppression of luteinizing hormone and follicle stimulating  
 CC hormone. This indirectly suppresses steriodogenesis and gamete maturation  
 CC in males and females. This type of anti-self response in humans is useful  
 CC in the treatment of prostate cancer and breast cancer. In livestock, the  
 CC ability to stimulate an anti-self response provides a simple alternative  
 CC to physical castration. Immunocastration of pigs is a better alternative  
 CC to physical castration, as it does not result in any of the detrimental  
 CC side effects associated with physical castration. Other examples of  
 CC diseases and conditions treated with self proteins fused with ubiquitin  
 CC are TNF and its epitopes to modulate septic shock, arthritis,  
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig  
 CC epsilon heavy chain for the control of allergic reactions; chorionic  
 CC gonadotropin for fertility control; and sperm proteins for fertility  
 CC control. A further use of the fusion proteins is as part of a vaccine to

CC enhance growth rate and thereby the final weight of the livestock prior  
 CC to shipment to market. In addition, the fusion proteins of the invention  
 CC can be used to detect and identify antibodies from experimental samples.  
 CC This sequence represents a GnRH mixed dimer used in the construction of  
 CC a ubiquitin fusion protein described in the method of the invention.

XX Sequence 20 AA;

Query Match 58.4%; Score 59; DB 20; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 0.14;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPR 11  
 |||||||:  
 Db 2 hwsyglrprg 11

RESULT 10

R07324 R07324 standard; protein; 21 AA.

AC R07324;

DT 29-JAN-1991 (first entry)

DE Lutetinsing hormone releasing hormone derived peptide.

XX LHRR; vaccine; meat; pigs; cancer; sterilisation.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /label=OTHER

FT Modified-site 3 /note="OTHER-pyroglyutamic acid

FT Modified-site 13 /label=OTHER

FT Modified-site 13 /note="OTHER- N-formyl-Trp (optional)"

FT Misc-difference 21 /label=OTHER

FT /note="OTHER- N-formyl-Trp (optional)"

XX WO9011298-A.

PN 04-OCT-1990.

XX 22-MAR-1990; 90WO-NL00037.

XX 23-MAR-1989; 89NL-0000726.

XX (DIER-) STICHT CENT DIERGEN.

XX Meloen RH, Wensing CUG;

XX WPI; 1990-320228/42.

PT Peptide for vaccinating mammals against LHRR - comprises at least  
 PT two lutetinsing hormone releasing hormone sequences in tandem

PS Claim 4; Page 10; 15pp; English.

XX The peptide comprises at least 2 LHRR sequences in tandem. The  
 CC peptide can be used to vaccinate mammals (e.g. pigs) against LHRR.  
 CC Such vaccination is used in human medicine for the treatment of  
 CC prostate cancer and breast cancer and some forms of hypophyseal  
 CC carcinoma. Other applications. Include sterilisation of domestic  
 CC animals and treatment of aggression in dogs. A major use of the  
 CC vaccination is to improve meat quality in pigs by avoiding "boar  
 CC odour" associated with the meat of sexually mature pigs.

CC See also R07323.

XX Sequence 21 AA;

Query Match 58.4%; Score 59; DB 11; Length 21;  
 Best Local Similarity 90.0%; Pred. No. 0.14;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPR 11  
 |||||||:  
 Db 2 hwsyglrprg 11

RESULT 11

R07323 R07323 standard; peptide; 30 AA.

AC R07323;

DT 29-JAN-1991 (first entry)

DE Lutetinsing hormone releasing hormone derived peptide.

XX LHRR; vaccine; meat; pigs; cancer; sterilisation.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /label=OTHER

FT /note="OTHER-pyroglyutamic acid or Gln having at  
 least one additional AA attached."

FT Modified-site 3 /label=OTHER

FT Modified-site 13 /note="OTHER- N-formyl-Trp (optional)"

FT Modified-site 13 /label=OTHER

FT Modified-site 10, 19 /note="OTHER- N-formyl-Trp (optional)"

FT Region /label=repeat

FT /note="repeat must occur at least once"

FT Misc-difference 30 /label=OTHER

FT /note="OTHER-Gly-NH2 or Gly having at  
 least one additional AA attached"

XX WO9011298-A.

PN 04-OCT-1990.

XX 22-MAR-1990; 90WO-NL00037.

XX 23-MAR-1989; 89NL-0000726.

XX (DIER-) STICHT CENT DIERGEN.

XX Meloen RH, Wensing CUG;

XX WPI; 1990-320228/42.

PT Peptide for vaccinating mammals against LHRR - comprises at least  
 PT two lutetinsing hormone releasing hormone sequences in tandem

PS Claim 2; Page 10; 15pp; English.

XX The peptide comprises at least 2 LHRR sequences in tandem. There  
 CC may be a spacer gp. between Gly(20) and Gln(21). The peptide can  
 CC be used to vaccinate mammals (e.g. pigs) against LHRR. Such  
 CC vaccination is used in human medicine for the treatment of prostate  
 CC cancer and breast cancer and some forms of hypophyseal carcinoma.  
 CC Other applications. Include sterilisation of domestic animals and  
 CC treatment of aggression in dogs. A major use of the vaccination is



CC / to improve meat quality in pigs by avoiding "boar odour" associated  
 CC with the meat of sexually mature pigs.  
 CC See also R07323.

XX Sequence 30 AA:

Query Match 58.4%; Score 59; DB 11; Length 30;  
 Best Local Similarity 90.0%; Pred. No. 0.19;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSYGLRPCR 11  
 |||||  
 Db 2 hwsyglrpgq 11

RESULT 12  
 Y31183  
 ID Y31183 standard; peptide: 40 AA.

AC Y31183;  
 XX  
 DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH fragment 2.

XX  
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;  
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;  
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;  
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;  
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;  
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;  
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;  
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.

XX Unidentified.

OS

PN W09942472-A1.

PD 26-AUG-1999.

PF 26-JAN-1999; 99WO-US01588.

PR 19-FEB-1998; 98US-0026276.

PA (IGEN-) IGEN INT INC.

PI Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;  
 WPI: 1999-518582/43.

PT Epitope-containing fusion proteins used to generate a highly  
 PT specific immune responses

PS Claim 83; Page 43; 67pp; English.

XX This invention describes a novel fusion protein, comprising a heat shock  
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner  
 CC which is useful for the stimulation of a highly specific immune response  
 CC when administered to an animal. The protein of the invention may be  
 CC post-translationally modified (e.g. by the addition of fatty acids to  
 CC enhance immunogenicity). The fusion proteins of the invention can be  
 CC used as vaccines to induce an immune response. When a T cell epitope is  
 CC attached, they can be used for control of viral infections, bacterial  
 CC infections, parasitic infection and cancer. The fusion proteins can be  
 CC used in pharmaceutical compositions for the treatment of gastrointestinal  
 CC diseases, pulmonary infections, respiratory infections, and HIV  
 CC infections. The use of ubiquitin as a scaffold is also useful for the  
 CC presentation and stimulation of anti-self immune responses, e.g.  
 CC generation of anti-gonadotropin releasing hormone antibodies which result  
 CC in the suppression of luteinizing hormone and follicle stimulating  
 CC hormone. This indirectly suppresses steriodogenesis and gamete maturation  
 CC in males and females. This type of anti-self response in humans is useful

CC in the treatment of prostate cancer and breast cancer. In livestock, the  
 CC ability to stimulate an anti-self response provides a simple alternative  
 CC to physical castration. Immunocastration of pigs is a better alternative  
 CC to physical castration, as it does not result in any of the detrimental  
 CC side effects associated with physical castration. Other examples of  
 CC diseases and conditions treated with self proteins fused with ubiquitin  
 CC are TNF and its epitopes to modulate septic shock, arthritis,  
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig  
 CC epsilon heavy chain for the control of allergic reactions; chorionic  
 CC gonadotropin for fertility control; and sperm proteins for fertility  
 CC control. A further use of the fusion proteins is as part of a vaccine to  
 CC enhance growth rate and thereby the final weight of the livestock prior  
 CC to shipment to market. In addition, the fusion proteins of the invention  
 CC can be used to detect and identify antibodies from experimental samples.  
 CC This sequence represents a GnRH fragment used in the construction of  
 CC a ubiquitin fusion protein described in the method of the invention.

XX Sequence 40 AA:

Query Match 58.4%; Score 59; DB 20; Length 40;  
 Best Local Similarity 90.0%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSYGLRPCR 11  
 |||||  
 Db 2 hwsyglrpgq 11

RESULT 13  
 Y31182  
 ID Y31182 standard; peptide: 41 AA.

AC Y31182;

DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH fragment.

XX  
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;  
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;  
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;  
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;  
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;  
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;  
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;  
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.

XX Unidentified.

OS

PN W09942472-A1.

PD 26-AUG-1999.

PF 26-JAN-1999; 99WO-US01588.

PR 19-FEB-1998; 98US-0026276.

PI Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;  
 WPI: 1999-518582/43.

PT Epitope-containing fusion proteins used to generate a highly  
 PT specific immune responses

PS Claim 81; Page 43; 67pp; English.

XX This invention describes a novel fusion protein, comprising a heat shock  
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner  
 CC which is useful for the stimulation of a highly specific immune response  
 CC when administered to an animal. The protein of the invention may be

CC post-translationally modified (e.g. by the addition of fatty acids to  
CC enhance immunogenicity). The fusion proteins of the invention can be  
CC used as vaccines to induce an immune response. When a T cell epitope is  
CC attached, they can be used for control of viral infections, bacterial  
CC infections, parasitic infections and cancer. The fusion proteins can be  
CC used in pharmaceutical compositions for the treatment of gastrointestinal  
CC diseases, pulmonary infections, respiratory infections, and HIV  
CC infections. The use of ubiquitin as a scaffold is also useful for the  
CC generation and stimulation of anti-self immune responses, e.g.  
CC generation of anti-gonadotropin releasing hormone antibodies which result  
CC in the suppression of luteinizing hormone and follicle stimulating  
CC hormone. This indirectly suppresses steroidogenesis and gamete maturation  
CC in males and females. This type of anti-self response in humans is useful  
CC in the treatment of prostate cancer and breast cancer. In livestock, the  
CC ability to stimulate an anti-self response provides a simple alternative  
CC to physical castration. Immunocastration of pigs is a better alternative  
CC to physical castration, as it does not result in any of the detrimental  
CC side effects associated with physical castration. Other examples of  
CC diseases and conditions treated with self proteins fused with ubiquitin  
CC are TME and its epitopes to modulate septic shock, arthritis,  
CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig  
CC epsilon heavy chain for the control of allergic reactions; chorionic  
CC gonadotropin for fertility control; and sperm proteins for fertility  
CC control. A further use of the fusion proteins is as part of a vaccine to  
CC enhance growth rate and thereby the final weight of the livestock prior  
CC to shipment to market. In addition, the fusion proteins of the invention  
CC can be used to detect and identify antibodies from experimental samples.  
CC This sequence represents a GnRH fragment used in the construction of  
CC a ubiquitin fusion protein described in the method of the invention.

XX Sequence 41 AA:

Query Match 58.4%; Score 59; DB 20; Length 41;  
Best Local Similarity 90.0%; Pred. No. 0.26; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 11  
| | | | | | | | | |  
Db 2 hwsyglrpg 11

RESULT 14  
W94891  
ID W94891 standard; peptide; 9 AA.  
XX  
AC W94891;

DT 11-MAY-1999 (first entry)

DE LHRR peptide fragment.

XX LHRR; immune response; luteinizing hormone releasing hormone; DT;  
KW diptheria toxoid; castrating; oestrus cycling; aggression; breast;  
KW sexual activity; organoleptic; livestock; cell growth; malignant;  
KW prostate; ovarian; oncofetal; hyperplastic; pregnancy;  
KW endometriosis; inflammatory response.

OS Homo sapiens.

PN W09902180-A1.

PD 21-JAN-1999.

PF 09-JUL-1998; 98MO-AUD0532.

PR 09-JUL-1997; 97AU-0007768.

PA (CSLC-) CSL LTD.

PI McNameara MK;

DR WPI; 1999-120511/10.

XX New immunogenic luteinizing hormone releasing hormone compositions  
PT comprise LHRR conjugated to diptheria toxoid and adsorbed to an  
PT ionic polysaccharide, used to inhibit reproductive function in  
PT animals

PS Example 3; Page 30; 41pp; English.

CC The invention relates immunogenic composition for eliciting an immune  
CC response to luteinizing hormone releasing hormone (LHRH). The composition  
CC comprises a LHRR-diptheria toxoid (DT) conjugate adsorbed to an ionic  
CC polysaccharide. The LHRR-DT compositions can be used for eliciting an  
CC immune response to LHRR, for castrating an animal, for regulating  
CC oestrus cycling in a female animal or for inhibiting characteristics  
CC induced by the sexual maturation of an animal, e.g. aggression or sexual  
CC activity. They can also be used for achieving production gains in  
CC livestock, e.g. reduction or elimination of unwanted organoleptic  
CC characteristics from the meat of livestock. They can also be used for  
CC inhibiting the growth of cells which are regulated directly or indirectly  
CC by LHRR, e.g. malignant breast cells, malignant prostate cells, malignant  
CC ovarian cells, malignant oncofetal cells or hyperplastic cells. They can  
CC also be used for down-regulating the libido of an animal. They can also  
CC be used for inhibiting pregnancy, prostate enlargement, endometriosis or  
CC inflammatory responses. The LHRR compositions induce a more effective  
CC immune response against LHRR than the LHRR-carrier adjuvant compositions.  
CC The effective immune response against LHRR results in prevention of the  
CC release of the hormones LH and FSH from the anterior pituitary. Sequences  
CC W94890-93 are peptide derivatives of LHRR.

SQ Sequence 9 AA:

Query Match 57.4%; Score 58; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.1e+05; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
| | | | | | | | | |  
Db 1 hwsyglrpg 9

RESULT 15  
P10097  
ID P10097 standard; peptide; 10 AA.  
XX  
AC P10097;

DT 19-AUG-1992 (first entry)

DE Sequence of luteinizing hormone (LH-RH, ICSH) liberating hormone.

XX Gonadorelin; luteinizing hormone releasing hormone; LH-RH;  
KW ICSH; prostatic hyperplasia therapy.

OS Mammal.

FH Key Location/Qualifiers

FT Misc-difference 1 /label= Pyr

FT Modified-site 10 /label= Gly-NH2

PN BE887639-A.

PD 24-AUG-1981.

PF 27-AUG-1981; 81BE-0303944.

PR 22-MAY-1980; 80US-0152241.

PA (AMHP) AVERST MCKENNA HARR.

PI Auctair C;

xx / WPI: 1981-66067D/37 (66067D).

DR Gonadorelin for treatment of benign prostatic hyperplasia - is  
 XX the deca-peptide Pyr-His-Tyr-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH2 or  
 PT luteinizing hormone liberating hormone  
 XX

PS Claim 1; Page 7; 9pp; French.

XX The inventors claim a compsn. for the reduct. or prevention of  
 CC undesired prostatic growth in males. The compsn. contains a  
 CC decapeptide (gonadorelin) (P10097) with an appropriate vehicle or  
 CC support. The compsn. is used for treating e.g. benign prostatic  
 CC hyperplasia by parenteral admin. in daily doses of 0.035-11.0 (pref.  
 CC 0.080-2.0) mg/Kg. Gonadorelin is the generic name for LH-RH and is  
 CC described in US3835108. In the example s.c. injection of the  
 CC compsn. significantly reduced the wt. of seminal vesicles and  
 CC ventral prostate in rats without affecting the wt. of the  
 CC testicles.  
 XX

SO Sequence 10 AA;

Query Match 57.4%; Score 58; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.099;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10  
 |||||  
 Db 2 hwsyglrpg 10

Search completed: March 2, 2001, 10:53:46  
 Job time: 911 sec

**THIS PAGE BLANK (USPTO)**

OM of: US-09-306-689-10 to: A\_Geneseq\_36:\* out\_format : ffs  
Date: Mar 2, 2001 10:31 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=framet-n2p.model -DEV=xlp  
-O=cg2\_1/USPTO.spool/US0930689/runat.02032001\_102824\_9522/app\_query.fasta.1.2389  
-DB=A\_Geneseq\_36 -OEMT=fastan -SUFFIX=rag -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfis  
-NORM=ext -MILEN=0 -MAXLEN=200000000  
-USER=US0930689\_@cg2\_1\_187 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPXY  
-WAIT -THREADS=1

## Search information block:

Query: US-09-306-689-10  
Query length: 147  
Database: A\_Geneseq\_36:\*  
Database sequences: 268485  
Search time (sec): 109.740000

## score\_list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation
/SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:W03944 +		230.00	646.90	4.4e-28	49
/SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:W79567 +		230.00	646.90	4.4e-28	49
/SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:W61542 +		230.00	646.90	4.4e-28	49
/SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:W58363 +		230.00	646.90	4.4e-28	49
/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y58363 +		230.00	646.90	4.4e-28	49
/SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:Y58135 +		230.00	646.90	4.4e-28	49
/SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:W03943 +		230.00	624.78	6.7e-28	544
/SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:W79570 +		230.00	624.78	6.7e-28	544
/SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:W79573 +		230.00	622.53	7.0e-28	695
/SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:Y58361 +		230.00	622.53	7.0e-28	695
/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y58133 +		230.00	622.53	7.0e-28	695
/SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:W03942 +		230.00	619.40	7.5e-28	977
/SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:W79569 +		230.00	619.40	7.5e-28	977
/SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:Y31183 +		199.50	448.79	5.7e-17	41
/SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:Y31187 +		187.50	403.33	2.4e-15	323
/SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:R11186 +		180.50	389.09	1.7e-14	283
/SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:W21650 +		170.50	384.13	2.1e-13	44
/SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:W21649 +		155.00	372.41	9.9e-13	42
/SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:R07323 +		135.00	353.43	1.6e-11	30
/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y79308 +		149.00	333.82	9.8e-11	60
/SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:Y31174 +		108.00	253.42	8.8e-06	20
/SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:Y31185 +		107.50	228.64	1.6e-05	263
/SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:W47438 +		105.00	246.80	2.1e-05	20
/SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:Y31178 +		105.00	246.80	2.1e-05	20
/SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:Y31179 +		105.00	246.80	2.1e-05	20
/SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:R07324 +		105.00	246.35	2.1e-05	21
/SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:Y31177 +		102.00	240.18	4.8e-05	20
/SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:W21648 +		91.00	215.45	0.0011	21
/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y89763 +		78.00	188.17	0.0423	18
/SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:R61427 +		76.50	169.87	0.0867	92
/SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:R11179 +		72.00	150.50	0.3722	257
/SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:R78282 +		71.00	173.25	0.3041	17
/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y8139 +		71.00	173.25	0.3041	17
/SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:R44956 +		70.00	171.60	0.3993	16
/SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:R11177 +		70.00	146.12	0.6551	256
/SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:R11178 +		70.00	146.08	0.6556	257
/SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:R11181 +		69.00	144.02	0.8677	253
/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y89759 +		68.50	151.63	0.8438	98
/SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:R75151 +		68.00	147.05	1.04	143

/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y54052 -	68.00	132.79	1.37	675
/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y54053 -	68.00	132.79	1.37	675
/SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y89760 +	67.50	163.15	0.8577	22
/SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:W89276 +	67.00	168.42	0.8730	11
/SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:R44953 +	67.00	166.20	0.9114	14

seq\_name: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:W03944

## seq\_documentation\_block:

ID W03944 standard; Protein; 49 AA.

AC W03944;

DT 20-NOV-1996 (first entry)

XX GnRH 4-repeat sequence.

DE Leukotoxin: LKT; gonadotropin-releasing hormone: GnRH;

KW fusion protein; immunogen; vaccine; fertility control;

KM contraceptive; sterilisation; pcB113; pcB111.

XX Synthetic.

OS Key

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

alignment\_scores:  
 Quality: 290.00 Length: 49  
 Ratio: 5.918 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-306-689-10 x W03944 ..

Align seg 1/1 to: W03944 from: 1 to: 49

```

1 CAGCATGGAGACTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 17
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34
51 CTACGGCTGCGTCCGGGTGACTAGCCAGCATTTGGAGCTAGCGGCTGC 100
|||||
101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGCGCTGCGCGGT 147
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

```

seq\_name: /SIDSI/gcgdata/geneseq/geneseq/A11998.DAT.W79567

seq\_documentation\_block:

ID W79567 standard; Protein: 49 AA.

AC W79567;

DT 24-DEC-1998 (first entry)

DE GnRH-2.

KM Gonadotropin releasing hormone; GnRH; decapeptide; hypothalamus; spacer;  
 LH; luteinising hormone; FSH; follicle stimulating hormone; vertebrate;  
 pyroGlu; chamera; leukotoxin polypeptide; multimer; vaccine; tumour;  
 immunogenic.

OS Synthetic.

PN W09806848-A1.

PD 19-FEB-1998.

PF 08-AUG-1997; 97WO-CA00559.

PR 09-AUG-1996; 96US-0694865.

PA (UYSA-) UNIV SASKATCHEWAN.

PI Manns JG, Potter AA;

DR WPI; 1998-159540/14.

DR N-PSDB; V61529.

Chimeric protein of leukotoxin and gonadotropin releasing hormone  
 useful for, e.g. preparation of vaccines for reduction of incidence  
 of mammary tumours in mammals

Disclosure: Figure 1B; 118pp; English.

The present sequence represents a recombinantly produced or chemically  
 synthesised gonadotropin releasing hormone-2 (GnRH-2) polypeptide, which  
 contains four copies of the GnRH decapeptide (V61529) and triplet amino  
 acid spacers between each of these sequences. This decapeptide is  
 secreted naturally by the hypothalamus which controls release of both the  
 luteinising hormone (LH) and the follicle stimulating hormone (FSH) in  
 vertebrates. This sequence, as compared to the native peptide, has been  
 found to have an N-terminal Gln rather than a pyroGlu residue, and also  
 contains substitutions at amino acid residues 15 and 41, whereby His is  
 replaced by Asp. This produces an alternating multimeric GnRH sequence  
 which is highly immunogenic that can be used in the construction of a  
 chimeric protein that comprises a leukotoxin polypeptide, several

CC multimers, and the GnRH sequence. The chimeric protein can be used as a  
 CC vaccine to help reduce the incidence of mammary tumours in a mammalian  
 CC individual.

XX Sequence 49 AA;

SD

alignment\_scores:  
 Quality: 290.00 Length: 49  
 Ratio: 5.918 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-306-689-10 x W79567 ..

Align seg 1/1 to: W79567 from: 1 to: 49

```

1 CAGCATTTGAGACTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 17
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34
51 CTACGGCTGCGTCCGGGTGACTAGCCAGCATTTGGAGCTAGCGGCTGC 100
|||||
101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGCGCTGCGCGGT 147
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

```

seq\_name: /SIDSI/gcgdata/geneseq/geneseq/A11998.DAT.W61542

seq\_documentation\_block:

ID W61542 standard; Protein: 49 AA.

AC W61542;

DT 27-OCT-1998 (first entry)

DE Peptide hormone GnRH-2 decapeptide (4 copies) fragment.

KM GnRH; gonadotropin releasing hormone peptide hormone; leukotoxin;  
 immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;  
 immune response; hormone receptor; cancerous cell; domestic animal;  
 porcine; bovine; luteinizing hormone; follicle stimulating hormone;  
 immunocastate.

OS Synthetic.

PN W09834639-A1.

PD 13-AUG-1998.

PF 04-FEB-1998; 98WO-CA00059.

PR 05-FEB-1997; 97US-0036883.

PA (BIOS-) BIOSTAR INC.

PI Acres SD, Harland R, Manns JG;

DR WPI; 1998-446952/38.

DR N-PSDB; V45190.

Immunisation against endogenous molecules by administering vaccine  
 to ear - useful to elicit efficient and uniform immune response  
 against e.g. gonadotropin releasing hormone to immunocastate pigs  
 and cattle

Example 1; Fig 1B; 61pp; English.

This represents the amino acid sequence of the gonadotropin releasing  
 hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-  
 GnRH polypeptide gene fusions. This is used to exemplify the method of

CC invention of immunisation against endogenous molecules by administering  
 CC a vaccine which comprises an immunogen and a carrier to the ear of the  
 CC mammal. The method is useful for eliciting an efficient and uniform  
 CC immune response to block or suppress the activity of an endogenous  
 CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,  
 CC or to elicit an immune response against a targeted endogenous cell type  
 CC (e.g. a cancerous or otherwise diseased cell). It is especially in porcine  
 CC to reduce the levels of GnRH in domestic animals, especially in porcine  
 CC or bovine species. The use of GnRH immunogens in the vaccine reduces the  
 CC levels of luteinizing hormone and follicle stimulating hormone and helps  
 CC in immunocastrating the animal. Administration of vaccine compositions to  
 CC the ear instead of intramuscular administration into the neck increases  
 CC the efficiency of vaccination of mammals against endogenous immunogens,  
 CC and may increase uniformity of vaccine presentation since the ear is  
 CC relatively uniform from animal to animal.

XX Sequence 49 AA;

#### alignment\_scores:

Quality: 290.00 Length: 49  
 Ratio: 5.918 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-09-306-689-10 x W61542 ..

Align seg 1/1 to: W61542 from: 1 to: 49

```

1 CAGCATTTGAGCTACGCGCTCGCCCTGGCAGCGGTTCATGATTGGAG 50
|||||
1 GlnH1StrpSerTyrGlyLeuArgProGlySerGlnH1StrpSerGlnAspTrpSe 17
51 CTACGGCCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGACTAGCGCCGTC 100
|||||
17 rTyrGlyLeuArgProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 34
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGCGCTCGCCGGGT 147
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.Y58363

#### seq\_documentation\_block:

ID Y58363 standard; Protein: 49 AA.

AC Y58363;

XX 27-MAR-2000 (first entry)

XX Four-copy gonadotropin-releasing hormone (GnRH) multimer.

XX GnRH multimer; gonadotropin-releasing hormone; immunosterilisation;

KW immunoccontraception; vaccine; feline; canine; equine; cervine; ds

OS Mammalia.

XX Synthetic.

XX WO962545-A2.

XX 09-DEC-1999.

XX 28-MAY-1999; 99WO-CA00493.

XX 04-JUN-1998; 98US-0088024.

XX 06-MAY-1999; 99US-0306689.

XX (BIOS-) BIOSTAR INC.

XX Robblins SC;

XX MPI; 2000-086857/07.

DR N-PSDB; 255702.

XX Hormone immunogens, analogues or antibodies used to manufacture  
 PT vaccines for suppression of reproductive behavior and fertility in  
 PT vertebrates -  
 XX Claim 5; Fig 5B; 88P; English.

CC This sequence represents a four-copy gonadotropin-releasing  
 CC hormone (GnRH) multimer, where the second and fourth GnRH sequence  
 CC have a His to Asp substitution at position 2 of the GnRH sequence.  
 CC The invention relates to GnRH immunogens, analogues or antibodies  
 CC that cross-react with endogenous GnRH of a vertebrate. A specifically  
 CC claimed immunogenic fusion protein (Y58361) comprises, in the N to  
 CC C-terminal direction, a synthetic peptide sequence (Y58364), an eight  
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer  
 CC sequence of Y58363), the LRT protein (which functions as a carrier  
 CC protein), and a second eight copy GnRH multimer. The fusion protein may  
 CC be used in a vaccine composition for prepubertal administration to a  
 CC vertebrate subject to result in prolonged suppression of reproductive  
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are  
 CC used to manufacture a composition or vaccine for immunosterilisation or  
 CC immunoccontraception of feline, canine, equine or cervine subjects.  
 CC The vaccines are used to suppress reproductive behaviour and/or  
 CC fertility for at least 10 months. The prepubertal administration  
 CC results in a prolonged, long-term suppression of testicular development  
 CC and/or function in males, or a prolonged, long-term suppression of  
 CC ovarian development and/or function in females. The methods provide a  
 CC viable and desirable alternative to surgical forms of sterilisation that  
 CC are currently used.

XX Sequence 49 AA;

#### alignment\_scores:

Quality: 290.00 Length: 49  
 Ratio: 5.918 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-09-306-689-10 x Y58363 ..

Align seg 1/1 to: Y58363 from: 1 to: 49

```

1 CAGCATTTGAGCTACGCGCTCGCCCTGGCAGCGGTTCATGATTGGAG 50
|||||
1 GlnH1StrpSerTyrGlyLeuArgProGlySerGlnH1StrpSerGlnAspTrpSe 17
51 CTACGGCCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGACTAGCGCCGTC 100
|||||
17 rTyrGlyLeuArgProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 34
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGCGCTCGCCGGGT 147
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.Y58135

#### seq\_documentation\_block:

ID Y58135 standard; Protein: 49 AA.

XX Y58135;

XX 07-MAR-2000 (first entry)

XX GnRH analogue multimer, containing four copies of the GnRH analogue.

XX Gonadotropin releasing hormone; GnRH; leukotoxin; LRT; fusion protein;

KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;

KW non-androgenic; steroid; reduction; weight gain; muscle distribution;

KW fat distribution; male pattern; boar taint; flavour; impairment;

KW reliable; immunocastration; meat production.

XX Synthetic.

```

OS Mammalia.
XX WO9956771-A2.
XX 11-NOV-1999.
XX
XX 05-MAY-1999; 99WO-CA00360.
XX
XX 05-MAY-1999; 98US-0084217.
XX
XX (BIOS-) BIOSTAR INC.
XX
XX Manns JG, Acres SD, Harland R;
XX
XX WPI; 2000-062125/05.
XX
XX N-PSDB; 246402.
XX
XX Production of uncastrated male food animals using vaccines -
XX
XX Example 1; Fig 2B; 87pp; English.
XX
XX This sequence represents four copies of a gonadotropin
XX releasing hormone (GNRH) analogue. DNA encoding which was
XX used in the construction of a chimeric GnRH-leukotoxin (LKT)
XX fusion gene (246400). This fusion gene encodes a GnRH-LKT fusion
XX protein which may be used as a vaccine. The LKT portion of the protein
XX acts to enhance the immunogenicity of the GnRH portion. The invention
XX relates to a method of using two GnRH immunogen vaccines to produce
XX uncastrated male animals for meat production, one vaccination prior to
XX or during the fattening period to reduce circulating testosterone levels,
XX and the second vaccination about 2-8 weeks before slaughter to
XX substantially reduce androgenic and/or non-androgenic steroids. The
XX invention is used to produce food animals that exhibit the weight gain
XX and muscle/fat distribution of male animals without the problems
XX associated with male animals. Such problems include "boar taint", a
XX urine-like odour found in cooked meat of uncastrated pigs which is
XX caused by steroids stored in the tissues, and similar flavour
XX impairments in the meat of other intact male animals. The invention is
XX more reliable than prior art immunocastration techniques.
XX
XX Sequence 49 AA:
XX
XX alignment_scores:
XX      Quality: 290.00      Length: 49
XX      Ratio: 5.918      Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 100.000
XX
XX alignment_block:
XX US-09-306-689-10 x Y58135 ..
XX
XX Align seq 1/1 to: Y58135 from: 1 to: 49
XX
1 CAGCATTTGAGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 17
51 CTACGGGCTCGTCCGGGTGGCTCTACGCAATTGGAGCTAGCGGCTGC 100
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34
101 GCCCTGGCAGCGGTAGCCAAAGATTGGAGCTACGGCTGCGCGGTT 147
34 rPrGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49
seq_name: /STD1/gcgdata/geneseq/geneseq/AA1996.DAT:W03943.
seq_documentation_block:
ID W03943 standard; Protein: 544 AA.
XX
XX W03943:
XX
XX 20-NOV-1996 (first entry)

```

```

XX DE LKT-GnRH protein fusion from PCB11.
XX
XX KW Leukotoxin; LKT: gonadotropin-releasing hormone; GnRH;
XX KW fusion protein; immunogen; vaccine; fertility control;
XX KW contraceptive; sterilisation.
XX
XX OS Chimeric Pasteurella haemolytica A1 strain B122;
XX OS Chimeric synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Domain 1..493
XX FT /Label= LKT
XX FT Domain 494..544
XX FT /Label= GnRH_repeat_domain
XX
XX PN W09624675-A1.
XX
XX PD 15-AUG-1996.
XX
XX PF 24-JAN-1996; 96WO-CA00049.
XX
XX PR 10-FEB-1995; 95US-0387156.
XX
XX PA (UYSA-) UNIT SASKATCHEWAN.
XX
XX PI Manns JG, Potter AA;
XX
XX WPI; 1996-384447/38.
XX
XX N-PSDB; T37177.
XX
XX Gonadotropin-releasing hormone multimer fusion proteins - with
XX PT leukotoxin polypeptide for increased immunogenicity, useful in
XX PT antifertility vaccine prodn.
XX
XX PS Claim 8; Fig 7A-7E; 87pp; English.
XX
XX CC A chimeric protein (W03943) is composed of a fusion between
XX CC a truncated leukotoxin (LKT-111) from Pasteurella haemolytica
XX CC and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat
XX CC sequence (see also W03944). It is the product of a chimeric
XX CC gene (T37177) produced by deleting an approx. 1300 bp sequence
XX CC from PCB113 (see also T37176) coding for amino acids 352-784
XX CC of LKT-352. Recombinant plasmid PCB111 (LKT 111:4 copy GnRH,
XX CC ATCC 69748) was obtd. Escherichia coli transformants produced
XX CC the chimeric protein, which is useful as a vaccine for fertility
XX CC control, esp. immunological sterilisation of domestic or farm
XX CC animals.
XX
XX SQ Sequence 544 AA:
XX
XX alignment_scores:
XX      Quality: 290.00      Length: 49
XX      Ratio: 5.918      Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 100.000
XX
XX alignment_block:
XX US-09-306-689-10 x W03943 ..
XX
XX Align seq 1/1 to: W03943 from: 1 to: 544
XX
1 CAGCATTTGAGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 510
494 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 510
51 CTACGGGCTCGTCCGGGTGGCTCTACGCAATTGGAGCTAGCGGCTGC 100
510 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 527
101 GCCCTGGCAGCGGTAGCCAAAGATTGGAGCTACGGCTGCGCGGTT 147
527 rPrGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 542

```



seq\_name: /SIDS1/gcdata/geneseq/geneseqp/AA1998.DAT:W79570

seq\_documentation\_block:

ID W79570 standard; Protein: 544 AA.

XX W79570;

DT 24-DEC-1998 (first entry)

DE LKT-GnRH chimeric protein.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone; multimer;

KW cytotoxic activity; antigen presentation; immune response; vaccine;

XX tumour.

OS Synthetic.

PN W09806848-A1.

PD 19-FEB-1998.

PF 08-AUG-1997; 97WO-CA00559.

PR 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1998-159540/14.

DR N-PSDB; V61532.

PT Chimeric protein of leukotoxin and gonadotropin releasing hormone  
PT useful for, e.g. preparation of vaccines for reduction of incidence  
PT of mammary tumours in mammals

XX PS Disclosure; Figure 7.1-5; 118pp; English.

XX CC The present sequence represents the LKT-GnRH chimeric protein from  
XX CC PCB11. This plasmid contains the LKT 111 polypeptide fused to  
XX CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity  
XX CC which enables there to be an increase in antigen presentation and thus an  
XX CC optimal immune response. The removal of this region also enables the  
XX CC truncated LKT to be expressed at much higher levels and allows the amount  
XX CC of antigen administered to be reduced. This chimeric protein comprises a  
XX CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
XX CC chimeric protein can be used as a vaccine to help reduce the incidence of  
XX CC mammary tumours in a mammalian individual.

SQ Sequence 544 AA;

alignment\_scores:

Quality: 290.00

Length: 49

Ratio: 5.918

Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x W79570 ..

Align seg 1/1 to: W79570 from: 1 to: 544

```
1 CAGCATGTGAGCTACGGCTCGCCCTGGACGGCTTCACAGATTGAG 50
|||||
494 GlnHisTSPserTyrGlyLeuArpProGlySerGlnAspTyrPse 510
|||||
51 CTACGGCCCTGCTCCGGGTGGCTTAGCCAGCATTTGAGACTAGCGCTGC 100
|||||
510 TyrTyrGlyLeuArpProGlyGlySerSerGlnHisTSPserTyrGlyLeuA 527
|||||
101 GCCCTGACGACGGTAGCCAAAGATTGGAGCTACGGCTCGCTCCGGGT 147
|||||
```

527 rPrProGlySerGlySerGlnAspTyrPseTyrGlyLeuArpProGly 542

seq\_name: /SIDS1/gcdata/geneseq/geneseqp/AA1998.DAT:W79573

seq\_documentation\_block:

ID W79573 standard; Protein: 695 AA.

XX W79573;

DT 24-DEC-1998 (first entry)

DE LKT-GnRH chimeric protein.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone; multimer;

KW cytotoxic activity; antigen presentation; immune response; vaccine;

XX tumour.

OS Synthetic.

PN W09806848-A1.

PD 19-FEB-1998.

PF 08-AUG-1997; 97WO-CA00559.

PR 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1998-159540/14.

DR N-PSDB; V61535

PT Chimeric protein of leukotoxin and gonadotropin releasing hormone  
PT useful for, e.g. preparation of vaccines for reduction of incidence  
PT of mammary tumours in mammals

XX PS Claim 9; Figure 9.1-6; 118pp; English.

XX CC The present sequence represents the LKT-GnRH chimeric protein from  
XX CC PCB122. This plasmid contains the LKT 111 polypeptide fused to sixteen  
XX CC copies of the GnRH peptide. In the pattern of: 8 copies of GnRH-LKT 111-8  
XX CC copies of GnRH. This chimera lacks cytotoxic activity which enables  
XX CC there to be an increase in antigen presentation and thus an optimal  
XX CC immune response. The removal of this region also enables the truncated  
XX CC LKT to be expressed at much higher levels and allows the amount of  
XX CC antigen administered to be reduced. This chimeric protein comprises a  
XX CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
XX CC chimeric protein can be used as a vaccine to help reduce the incidence of  
XX CC mammary tumours in a mammalian individual.

SQ Sequence 695 AA;

alignment\_scores:

Quality: 290.00

Length: 49

Ratio: 5.918

Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x W79573 ..

Align seg 1/1 to: W79573 from: 1 to: 695

```
1 CAGCATGTGAGCTACGGCTCGCCCTGGACGGCTTCACAGATTGAG 50
|||||
9 GlnHisTSPserTyrGlyLeuArpProGlySerGlySerGlnAspTyrPse 25
|||||
51 CTACGGCCCTGCTCCGGGTGGCTTAGCCAGCATTTGAGACTAGCGCTGC 100
|||||
25 TyrTyrGlyLeuArpProGlyGlySerSerGlnHisTSPserTyrGlyLeuA 42
|||||
```

101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCGCGGCT 147  
 ||||||||||||||||||||||||||||||||||||||||||||  
 42 rprGlySerGlySerGlnAspTrpSerTrpGlyLeuArgProGly 57

seq\_name: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.Y58361

seq\_documentation\_block:  
 ID Y58361 standard; Protein: 695 AA.

XX Y58361;

XX 27-MAR-2000 (first entry)

XX Leukotoxin/gonadotropin-releasing hormone fusion protein.

XX Leukotoxin; gonadotropin-releasing hormone; GnRH; immunosterilisation;  
 KW immunoreception; vaccine; feline; canine; equine; cervine.

XX Chimeric - Pasteurella haemolytica.

OS Chimeric - Mammalia.

XX WO9962545-A2.

XX 09-DEC-1999.

XX 28-MAY-1999; 99WO-CA00493.

XX 04-JUN-1998; 98US-0088024.

XX 06-MAY-1999; 99US-0306689.

XX (BIOS-) BIOSTAR INC.

XX Robins SC;

XX WPI; 2000-086857/07.

XX N-PSDB; Z55700.

XX Hormone immunogens, analogues or antibodies used to manufacture  
 PT vaccines for suppression of reproductive behavior and fertility in  
 PT vertebrates -

PS Claim 20; Fig 6A-6F; 88pp; English.

XX This sequence represents a fusion protein comprising gonadotropin-  
 CC releasing hormone (GnRH) immunogens and a Pasteurella  
 CC haemolytica leukotoxin (LKT) protein. The fusion protein comprises, in  
 CC the N to C-terminal direction, a synthetic peptide sequence (Y58364), an  
 CC eight copy GnRH multimer (composed of two copies of the 4xGnRH multimer  
 CC sequence of Y58363), the LKT protein (which functions as a carrier  
 CC protein), and a second eight copy GnRH multimer. The fusion protein may  
 CC be used in a vaccine composition for prepubertal administration to a  
 CC vertebrate subject to result in prolonged suppression of reproductive  
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies  
 CC that cross-react with endogenous GnRH of a vertebrate subject are used  
 CC to manufacture a composition or vaccine for immunosterilisation or  
 CC immunoreception of feline, canine, equine or cervine subjects.  
 CC The vaccines are used to suppress reproductive behaviour and/or  
 CC fertility for at least 10 months. The prepubertal administration  
 CC results in a prolonged, long-term suppression of testicular development  
 CC and/or function in males, or a prolonged, long-term suppression of  
 CC ovarian development and/or function in females. The methods provide a  
 CC viable and desirable alternative to surgical forms of sterilisation that  
 CC are currently used.

XX Sequence 695 AA;

alignment\_scores:

Quality: 290.00

Length: 49

Ratio: 5.918 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x Y58361 ..  
 Align seq 1/1 to: Y58361 from: 1 to: 695

1 CAGCATTTGGAGCTACGGCTGCGCGGTAGCCAGATTGAGCTACGGCTGC 50  
 ||||||||||||||||||||||||||||||||||||||||||||  
 9 GlnHstPrpSerTrpGlyLeuArgProGlySerGlnHstPrpSerTrpGlyLeu 25  
 51 CTACGGCTGCGCGGTAGCCAGATTGAGCTACGGCTGC 100  
 ||||||||||||||||||||||||||||||||||||||||||||  
 25 rTyGlyLeuArgProGlySerGlnHstPrpSerTrpGlyLeu 42  
 101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCGCGGCT 147  
 ||||||||||||||||||||||||||||||||||||||||||||  
 42 rprGlySerGlySerGlnAspTrpSerTrpGlyLeuArgProGly 57

seq\_name: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.Y58133

seq\_documentation\_block:  
 ID Y58133 standard; Protein: 695 AA.

XX Y58133;

XX 07-MAR-2000 (first entry)

XX Gonadotropin releasing hormone-leukotoxin fusion protein.

XX Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;  
 KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;  
 KW non-androgenic; steroid; reduction; weight gain; muscle distribution;  
 KW fat distribution; male pattern; boar taint; flavour; impairment;  
 KW reliable; immunocastration; meat production.

XX Chimeric - Mammalia.

XX Chimeric - Pasteurella haemolytica.

XX WO9956771-A2.

XX 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00360.

XX 05-MAY-1998; 98US-0084217.

XX (BIOS-) BIOSTAR INC.

XX Manns JG, Acres SD, Harland R;

XX WPI; 2000-062125/05.

XX N-PSDB; Z46400.

XX Production of uncastrated male food animals using vaccines -

PS Claim 22; Fig 3A-3F; 87pp; English.

XX This sequence represents a chimeric gonadotropin releasing  
 CC hormone (GnRH)-leukotoxin (LKT) fusion protein, which may be  
 CC used as a vaccine. The LKT portion of the protein acts to enhance  
 CC the immunogenicity of the multimeric GnRH portion (Y58135). The invention  
 CC relates to a method of using two GnRH immunogen vaccines to produce  
 CC uncastrated male animals for meat production, one vaccination prior to  
 CC or during the fattening period to reduce circulating testosterone levels,  
 CC and the second vaccination about 2-8 weeks before slaughter to  
 CC substantially reduce androgenic and/or non-androgenic steroids. The  
 CC invention is used to produce food animals that exhibit the weight gain  
 CC and muscle/fat distribution of male animals without the problems  
 CC associated with male animals. Such problems include "boar taint", a  
 CC urine-like odour found in cooked meat of uncastrated pigs which is  
 CC caused by steroids stored in the tissues, and similar flavour  
 CC impairments in the meat of other intact male animals. The invention is  
 CC more reliable than prior art immunocastration techniques.

XX Sequence 695 AA;

## alignment\_scores:

Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x Y58133 ..

Align seg 1/1 to: Y58133 from: 1 to: 695

```

1 CAGCATTGAGCTACGGCCCTGCGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
9 GlnHistrpserTyrGlyLeuArgProGlySerGlySerGlnAspTrpse 25
25 rTyrglyLeuArgProGlyGlySerSerGlnHistrpserTyrGlyLeuA 42
51 CTACGGCCTGCGTCCGGGTGCTTAGCCAGCATTTGGAGCTACGGCCTGC 100
|||||
101 GCCCTGGCAGCGGTACGCCAAGATTGGAGCTACGGCCTCGCGGT 147
|||||
42 rTyrglyLeuArgProGlyGlySerSerGlnHistrpserTyrGlyLeuA 57
seq_name: /SIDSI/gcgsdata/geneseq/geneseq/AA1996.DAT:W03942

```

## seq\_documentation\_block:

ID W03942 standard; Protein: 977 AA.

AC W03942:

DT 20-NOV-1996 (first entry)

DE LKT-GnRH protein fusion from PCB113.

DE Leukotoxin: LKT; gonadotropin-releasing hormone: GnRH;

KM fusion protein; immunogen; vaccine; fertility control;

KM contraceptive; sterilisation.

OS Chimeric Pasteurella haemolytica A1 strain B122;

OS Chimeric synthetic.

Key Location/Qualifiers

FF Domain

FF Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

CC chimeric gene (T37176) produced by ligating a synthetic sequence  
CC for the 4-copy GnRH into vector pAA352 (ATCC 66283), which carries  
CC the LKT-352 gene. Recombinant plasmid PCB113 (LKT 352:4 copy  
CC GnRH, ATCC 69749) was obtained. Escherichia coli transformants  
CC produced the chimeric protein, which is useful as a vaccine for  
CC fertility control, esp. immunological sterilisation of  
CC domestic or farm animals.

SQ Sequence 977 AA:

alignment\_scores:  
Quality: 290.00 Length: 49  
Ratio: 5.918 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-306-689-10 x W03942 ..

Align seg 1/1 to: W03942 from: 1 to: 977

```

1 CAGCATTGAGCTACGGCCCTGCGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
927 GlnHistrpserTyrGlyLeuArgProGlySerGlySerGlnAspTrpse 943
51 CTACGGCCTGCGTCCGGGTGCTTAGCCAGCATTTGGAGCTACGGCCTGC 100
|||||
101 GCCCTGGCAGCGGTACGCCAAGATTGGAGCTACGGCCTCGCGGT 147
|||||
943 rTyrglyLeuArgProGlyGlySerSerGlnHistrpserTyrGlyLeuA 960
960 rTyrglyLeuArgProGlyGlySerSerGlnHistrpserTyrGlyLeuA 975
seq_name: /SIDSI/gcgsdata/geneseq/geneseq/AA1998.DAT:W79569

```

## seq\_documentation\_block:

ID W79569 standard; Protein: 977 AA.

AC W79569:

DT 24-DEC-1998 (first entry)

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

DE LKT-GnRH chimeric protein.

Dislosure: Figure 5.1-8; 118pp; English.

The present sequence represents the LKT-GnRH chimeric protein from  
CC PCB113. This plasmid contains the LKT 352 polypeptide (W79568) fused to

CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity  
 CC which enables there to be an increase in antigen presentation and thus an  
 CC optimal immune response. The removal of this region also enables the  
 CC truncated LRT to be expressed at much higher levels and allows the amount  
 CC of antigen administered to be reduced. This chimeric protein comprises a  
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The  
 CC chimeric protein can be used as a vaccine to help reduce the incidence of  
 CC mammary tumours in a mammalian individual.

XX Sequence 977 AA:

alignment\_scores:

Quality: 290.00 Length: 49  
 Ratio: 5.918 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-306-689-10 x W79569 ..

Align seg 1/1 to: W79569 from: 1 to: 977

1 CAGATTGGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50  
 ||||||||||||||||||||||||||||||||||||||||||||  
 927 GlnHsrpSerTyrglyLeuAtrProglySerGlySerGlnAspTrpse 943  
 51 CTACGGCGCTGCGGCTGCTAGCCAGCATTTGGAGCTACGGCTGC 100  
 ||||||||||||||||||||||||||||||||||||||||||||  
 943 rTyrglyLeuAtrProglySerGlySerGlnHsrpSerTyrglyLeuA 960  
 101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCTGCGGCT 147  
 ||||||||||||||||||||||||||||||||||||||||||||  
 960 rGProglySerGlySerGlnAspTrpSerTyrglyLeuAtrProgly 975

seq\_name: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.Y31183

seq\_documentation\_block:

ID Y31183 standard; peptide: 40 AA.

XX Y31183;

DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH fragment 2.

XX Ubiquitin: immunocastration; fusion protein; heat shock protein; epitope;  
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;  
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;  
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;  
 KW steroidogenesis; gamete maturation; prostate; breast; castration; TNF;  
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;  
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;  
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.

XX unidentified.

XX OS

XX MO942472-AL.

XX 26-AUG-1999.

XX 26-JAN-1999; 99MO-US01588.

XX 19-FEB-1998; 98US-0026276.

XX (IGEN-) IGEN INT INC.

XX Kenton JH, Lohmas GT, Pilon AL, Roberts SF, Tramontano A;

XX WPI; 1999-518582/43.

PT Epitope-containing fusion proteins used to generate a highly  
 PT specific immune responses  
 XX

PS Claim 83; Page 43; 67p; English.

XX This invention describes a novel fusion protein, comprising a heat shock  
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner  
 CC which is useful for the stimulation of a highly specific immune response  
 CC when administered to an animal. The protein of the invention may be  
 CC post-translationally modified (e.g. by the addition of fatty acids to  
 CC enhance immunogenicity). The fusion proteins of the invention can be  
 CC used as vaccines to induce an immune response. When a T cell epitope is  
 CC attached, they can be used for control of viral infections, bacterial  
 CC infections, parasitic infection and cancer. The fusion proteins can be  
 CC used in pharmaceutical compositions for the treatment of gastrointestinal  
 CC diseases, pulmonary infections, respiratory infections, and HIV  
 CC infections. The use of ubiquitin as a scaffold is also useful for the  
 CC presentation and stimulation of anti-self immune responses, e.g.  
 CC generation of anti-gonadotropin releasing hormone antibodies which result  
 CC in the suppression of luteinizing hormone and follicle stimulating  
 CC hormone. This indirectly suppresses steroidogenesis and gamete maturation  
 CC in males and females. This type of anti-self response in humans is useful  
 CC in the treatment of prostate cancer and breast cancer. In livestock, the  
 CC ability to stimulate an anti-self response provides a simple alternative  
 CC to physical castration. Immunocastration of pigs is a better alternative  
 CC to physical castration, as it does not result in any of the detrimental  
 CC side effects associated with physical castration. Other examples of  
 CC diseases and conditions treated with self proteins fused with ubiquitin  
 CC are TNF and its epitopes to modulate septic shock, arthritis,  
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig  
 CC epsilon heavy chain for the control of allergic reactions; chorionic  
 CC gonadotropin for fertility control; and sperm proteins for fertility  
 CC control. A further use of the fusion proteins is as part of a vaccine to  
 CC enhance growth rate and thereby the final weight of the livestock prior  
 CC to shipment to market. In addition, the fusion proteins of the invention  
 CC can be used to detect and identify antibodies from experimental samples.  
 CC This sequence represents a GnRH fragment used in the construction of  
 CC a ubiquitin fusion protein described in the method of the invention.

XX Sequence 40 AA:

alignment\_scores:

Quality: 199.50 Length: 49  
 Ratio: 5.250 Gaps: 3  
 Percent Similarity: 77.551 Percent Identity: 77.551

alignment\_block:

US-09-306-689-10 x Y31183 ..

Align seg 1/1 to: Y31183 from: 1 to: 40

1 CAGATTGGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50  
 ||||||||||||||||||||||||||||||||||||||||||||  
 1 GlnHsrpSerTyrglyLeuAtrProgly.....GlnHsrpse 14  
 51 CTACGGCGCTGCGGCTGCTAGCCAGCATTTGGAGCTACGGCTGC 100  
 ||||||||||||||||||||||||||||||||||||||||||||  
 14 rTyrglyLeuAtrProgly.....GlnHsrpSerTyrglyLeuA 28  
 101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCTGCGGCT 147  
 ||||||||||||||||||||||||||||||||||||||||||||  
 28 rGProgly.....GlnHsrpSerTyrglyLeuAtrProgly 40

seq\_name: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.Y31182

seq\_documentation\_block:

ID Y31182 standard; peptide: 41 AA.

XX Y31182;

DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH fragment.

XX Ubiquitin: immunocastration; fusion protein; heat shock protein; epitope;  
 KW

KM immune response stimulation; vaccine; T cell; viral; infection; cancer;  
 KM bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;  
 KM pulmonary infection; respiratory infection; scaffold; anti-self; pig;  
 KM steriodogenesis; gamete maturation; prostate; breast; castration; TNF;  
 KM tumour necrosis factor; septic shock; arthritis; Crohn's disease;  
 KM inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;  
 KM fertility; sperm protein; growth rate; antibody; detection; GnRH.  
 XX unidentified.  
 OS  
 XX  
 PN MO9942472-A1.  
 XX  
 PD 26-AUG-1999.  
 XX  
 PE 26-JAN-1999; 99WO-US01588.  
 XX  
 PR 19-FEB-1998; 98US-0026276.  
 XX  
 PA (IGEN-) IGEN INT INC.  
 XX  
 PI Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;  
 DR WPI: 1999-518582/43.  
 XX  
 PT Epitope-containing fusion proteins used to generate a highly  
 PS specific immune responses  
 XX  
 PS Claim 81; Page 43; 67pp; English.  
 XX  
 CC This invention describes a novel fusion protein, comprising a heat shock  
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner  
 CC which is useful for the stimulation of a highly specific immune response  
 CC when administered to an animal. The protein of the invention may be  
 CC post-translationally modified (e.g. by the addition of fatty acids to  
 CC enhance immunogenicity). The fusion proteins of the invention can be  
 CC used as vaccines to induce an immune response. When a T cell epitope is  
 CC attached, they can be used for control of viral infections, bacterial  
 CC infections, parasitic infection and cancer. The fusion proteins can be  
 CC used in pharmaceutical compositions for the treatment of gastrointestinal  
 CC diseases, pulmonary infections, respiratory infections, and HIV  
 CC infections. The use of ubiquitin as a scaffold is also useful for the  
 CC presentation and stimulation of anti-self immune responses, e.g.  
 CC generation of anti-gonadotropin releasing hormone antibodies which result  
 CC in the suppression of luteinizing hormone and follicle stimulating  
 CC hormone. This indirectly suppresses steriodogenesis and gamete maturation  
 CC in males and females. This type of anti-self response in humans is useful  
 CC in the treatment of prostate cancer and breast cancer. In livestock, the  
 CC ability to stimulate an anti-self response provides a simple alternative  
 CC to physical castration. Immunocastration of pigs is a better alternative  
 CC to physical castration, as it does not result in any of the detrimental  
 CC side effects associated with physical castration. Other examples of  
 CC diseases and conditions treated with self proteins fused with ubiquitin  
 CC are TNF and its epitopes to modulate septic shock, arthritis,  
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig  
 CC epsilon heavy chain for the control of allergic reactions; chorionic  
 CC gonadotropin for fertility control; and sperm proteins for fertility  
 CC control. A further use of the fusion proteins is as part of a vaccine to  
 CC enhance growth rate and thereby the final weight of the livestock prior  
 CC to shipment to market. In addition, the fusion proteins of the invention  
 CC can be used to detect and identify antibodies from experimental samples.  
 CC This sequence represents a GnRH fragment used in the construction of  
 CC a ubiquitin fusion protein described in the method of the invention.  
 CC  
 XX  
 SQ Sequence 41 AA:

alignment\_scores:                      length: 49  
                   quality: 199.50  
                   ratio: 5.250  
 Percent Similarity: 77.551            caps: 3  
    Percent Identity: 77.551

alignment\_block:  
 US-09-306-689-10 x Y31182

Align seg 1/1 to: Y31182 from: 1 to: 41  
 1 CAGCATTTGAGCTACGCGCTGCGCCCTGGCAGCGGTTCACAGATTGAG 50  
 1 GlnHisTrpSerTyrGlyLeuArgProGly.....GlnHisTrp 14  
 51 CTACGGCGCTGCGTCCGGGTGGCTCTAGCCAGCATTTGGAGCTACGCGCTC 100  
 14 rTyrGlyLeuArgProGly.....GlnHisTrpSerTyrGlyLeuA 28  
 101 GCCCTGCGCAGCGGTACGCAAGATTGGAGCTACGCGCTGCGCGGT 147  
 28 rGProGly.....GlnHisTrpSerTyrGlyLeuArgProGly 40  
 seq\_name: /SIDSL/gcdata/geneseq/geneseq/AA1991.DAT:R11187  
 seq\_documentation\_block:  
 ID R1187 standard; Protein; 323 AA.  
 AC R1187;  
 XX  
 DT 22-MAY-1991 (first entry)  
 XX  
 DE Plasmid pBTa859-encoded Tratp-multiple LHRH analogue fusion.  
 XX  
 KW Tratp protein; Luteinizing hormone releasing hormone; fusion protein;  
 KW immunological castration.  
 XX  
 FH Key location/Qualifiers  
 FT Peptide 1..20  
 FT /label= Tratp signal  
 FT Peptide 201..280  
 FT /label= 8 LHRH analogues in tandem repeat  
 XX  
 PN MO9102799-A.  
 FT 07-MAR-1991.  
 XX  
 PD 24-AUG-1990; 90WO-AU00373.  
 XX  
 PE 25-AUG-1989; 89AU-0005979.  
 PR (BIOT-) BIOTECHN AUST PTY L.  
 PA Russell-Jones GJ, Stewart AG, Tsonis CG;  
 PI WPI: 1991-087282/12.  
 DR N-PSDB: Q11021.  
 XX  
 PT Fusion proteins comprising LHRH analogue and Tratp (analogue) -  
 PT useful in vaccine for inhibition or control of reproduction in  
 PT vertebrates, esp. domestic animals  
 PS  
 PS Example 1: Fig 2 and 5; 53pp; English.  
 CC Plasmid pBTa859 is a Tratp-LHRH analogue fusion in which 8 copies  
 CC of an LHRH analogue have been inserted between amino acids 200 and  
 CC 201 of Tratp (Ogata R.T. et al., (1982) J.Bacteriol. 151:819-827).  
 CC The plasmid was constructed by two successive additions of DNA  
 CC coding for a dimer of LHRH analogue into the SmaI site of pBTa862  
 CC (see Q11020) which all ready carries four copies of the LHRH  
 CC sequence. After transformation, colonies with 8 LHRH molecules were  
 CC identified. Fusion proteins with multiple inserts generated a higher  
 CC anti-LHRH response (as measured by the binding of (125)I-LHRH at a  
 CC serum dilution of 1:2000 final) than constructs with a single  
 CC insert, in outbred mice and dogs. The fusion proteins can be used to  
 CC inhibit reproductive functions in vertebrates.  
 CC See also Q10995, Q10997-Q11000, Q11014-Q11020.  
 XX  
 SQ Sequence 323 AA:

alignment\_scores:                   Quality: 187.50                   Length: 49  
                                   Ratio: 4.934                   Gaps: 3  
 Percent Similarity: 77.551           Percent Identity: 69.388

alignment\_block:  
 US-09-306-689-10 x R1187   ..

Align seg 1/1 to: R1187 from: 1 to: 323

```

1 CAGCATTGGAGCTACGGCCTGCGCCCTGGCAGCGGTTCTCAGATTGGAG 50
   :::::::::::::::::::::::::::::::::::::  ::::  |||||
201 GluHistrpserTyrglyLeuArgProGly.....GluHistrpSe 214
51 CTACGGCCTGCGCTCCGGGTGCTCTAGCCAGCATTGGAGCTAGCGCCTGC 100
   ||:::::::::::::::::::::::::::  ::::::::::::::::::::::
214 rTyrglyLeuArgProGly.....GluHistrpserTyrglyLeuA 228
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGGCTGCGCTCCGGGT 147
   ||::::::::::::  ::::  ||::::::::::::::::::::::::::
228 rgProGly.....GluHistrpserTyrglyLeuArgProGly 240
  
```